THE UNIVERSITY OF WARWICK

Library Declaration and Deposit Agreement

1. STUDENT DETAILS

Please complete the following:

Full name: ……

University ID number: ……

2. THESIS DEPOSIT

2.1 I understand that under my registration at the University, I am required to deposit my thesis with the University in BOTH hard copy and in digital format. The digital version should normally be saved as a single pdf file.

2.2 The hard copy will be housed in the University Library. The digital version will be deposited in the University's Institutional Repository (WRAP). Unless otherwise indicated (see 2.3 below) this will be made openly accessible on the Internet and will be supplied to the British Library to be made available online via its Electronic Theses Online Service (EThOS) service.
[At present, theses submitted for a Master's degree by Research (MA, MSc, LLM, MS or MMedSci) are not being deposited in WRAP and not being made available via EthOS. This may change in future.]

2.3 In exceptional circumstances, the Chair of the Board of Graduate Studies may grant permission for an embargo to be placed on public access to the hard copy thesis for a limited period. It is also possible to apply separately for an embargo on the digital version. (Further information is available in the *Guide to Examinations for Higher Degrees by Research*.)

2.4 *If you are depositing a thesis for a Master's degree by Research, please complete section (a) below. For all other research degrees, please complete both sections (a) and (b) below:*

(a) Hard Copy

I hereby deposit a hard copy of my thesis in the University Library to be made publicly available to readers (please delete as appropriate) EITHER immediately OR after an embargo period of ……..... months/years as agreed by the Chair of the Board of Graduate Studies.

I agree that my thesis may be photocopied. YES / NO (*Please delete as appropriate*)

(b) Digital Copy

I hereby deposit a digital copy of my thesis to be held in WRAP and made available via EThOS.

Please choose one of the following options:

EITHER My thesis can be made publicly available online. YES / NO (*Please delete as appropriate*)

OR My thesis can be made publicly available only after…..[date] (Please give date)
YES / NO (*Please delete as appropriate*)

OR My full thesis cannot be made publicly available online but I am submitting a separately identified additional, abridged version that can be made available online.
YES / NO (*Please delete as appropriate*)

OR My thesis cannot be made publicly available online. YES / NO (*Please delete as appropriate*)

3. GRANTING OF NON-EXCLUSIVE RIGHTS

Whether I deposit my Work personally or through an assistant or other agent, I agree to the following:

Rights granted to the University of Warwick and the British Library and the user of the thesis through this agreement are non-exclusive. I retain all rights in the thesis in its present version or future versions. I agree that the institutional repository administrators and the British Library or their agents may, without changing content, digitise and migrate the thesis to any medium or format for the purpose of future preservation and accessibility.

4. DECLARATIONS

(a) I DECLARE THAT:

- I am the author and owner of the copyright in the thesis and/or I have the authority of the authors and owners of the copyright in the thesis to make this agreement. Reproduction of any part of this thesis for teaching or in academic or other forms of publication is subject to the normal limitations on the use of copyrighted materials and to the proper and full acknowledgement of its source.

- The digital version of the thesis I am supplying is the same version as the final, hard-bound copy submitted in completion of my degree, once any minor corrections have been completed.

- I have exercised reasonable care to ensure that the thesis is original, and does not to the best of my knowledge break any UK law or other Intellectual Property Right, or contain any confidential material.

- I understand that, through the medium of the Internet, files will be available to automated agents, and may be searched and copied by, for example, text mining and plagiarism detection software.

(b) IF I HAVE AGREED (in Section 2 above) TO MAKE MY THESIS PUBLICLY AVAILABLE DIGITALLY, I ALSO DECLARE THAT:

- I grant the University of Warwick and the British Library a licence to make available on the Internet the thesis in digitised format through the Institutional Repository and through the British Library via the EThOS service.

- If my thesis does include any substantial subsidiary material owned by third-party copyright holders, I have sought and obtained permission to include it in any version of my thesis available in digital format and that this permission encompasses the rights that I have granted to the University of Warwick and to the British Library.

5. LEGAL INFRINGEMENTS

I understand that neither the University of Warwick nor the British Library have any obligation to take legal action on behalf of myself, or other rights holders, in the event of infringement of intellectual property rights, breach of contract or of any other right, in the thesis.

Please sign this agreement and return it to the Graduate School Office when you submit your thesis.

Student's signature: .....…… Date: .....

Mechanisms Underlying Epigenetic Gene Silencing in Maize

David Gerald Schafer

A thesis submitted in fulfilment of the

requirements for the degree of

Doctor of Philosophy in Plant and Environmental Sciences

University of Warwick, Department of Life Sciences

May 2013

'Everything should be made as simple as possible, but no simpler'

-Albert Einstein

Contents

Abstract

Higher organisms can regulate gene expression through changes in epigenetic marks present on the genome. However, how this regulation takes place in organisms with highly repetitive/complex genomes is not well understood.

The acquisition of *de novo* DNA methylation in plants is mediated by siRNAs through the RNA-directed DNA methylation (RdDM) pathway. The targeted deposition of DNA methylation by this pathway allows for the transcriptional silencing of transposable elements and repeat sequences within the genome, as well as regulating gene expression. In addition, it has been hypothesized that mobile siRNAs may be involved in the epigenetic communication between different seed components. Thus the mobility of non-coding RNAs from extra-embryonic tissues could contribute to epigenetic modifications that could be transmitted to the offspring.

The aim of my thesis is to characterise the mechanisms involved in epigenetic gene silencing in maize through the use of a novel transgenic reporter. My work has identified components of the RdDM pathway to be involved in maintenance of gene silencing and show that imprinting and paramutation could be recapitulated using synthetic transgenes. In addition, I developed a novel grafting technique to demonstrate that epigenetic gene silencing could be efficiently transmitted between different seed components.

Collectively, this work provides an insight into the complex mechanisms that regulate gene expression in the highly repetitive/complex genome of maize.

Acknowledgments

I would like to first thank my supervisor, Dr Jose Gutierrez-Marcos for the invaluable advice and encouragement over the course of my PhD project. I would also like to thank Dr. Giorgio Perrella (University of Glasgow) and Dr. Simon Engledow for all their help along the way. Furthermore, I would like to thank everyone else in Marcos group, past and present, for their support and assistance when I have needed it and to everyone who has provided me with seed stocks, technical advice or lab chemicals. Finally a special thanks to friends and family who have kept me sane over the past 4 years, which include the Batemans, Charlie, Anna, Matt, Fresh, Andy and Gilly. This project could not have been completed without them.

Declaration

The work presented in this thesis is original, and has not been published or presented for any other degree.

List of Figures

List of Tables

List of Abbreviations

a	active
bp	base pair
cDNA	complementary Deoxyribonucleic Acid
°C	degrees Centigrade
cm	centimetres
CYR	Cell wall YFP Reporter
BSA	Bovine Serum Albumin
BTG	b1 genomic Transgene
DAP	Days After Pollination
DEPC	Diethylpyrocarbonate
DNA	Deoxyribonucleic Acid
dNTPs	Deoxynucleotide Triphosphates
ds	double-stranded
DTT	Dithiothreitol
EDTA	Ethylenediaminetetra-acetic Acid
F1	first filial generation
Fig.	Figure
µg	micrograms
g	grams
gDNA	genomic Deoxyribonucleic Acid
GFP	Green Fluorescent Protein
I1	first introgression generation
IR	Inverted Repeat
l	litres
LB	Luria Broth
µl	microlitres
µm	micrometres
µM	micromolar
mM	millimolar
M	Molar
MBD	Methyl-Binding Domains
min	minutes

ml	millilitres
mg	milligrams
mRNA	messenger Ribonucleic Acid
MS	Murashige and Skroog medium
ng	nanograms
nm	nanometres
nt	nucleotidesT
NYR	Nuclear YFP Reporter
PCR	Polymerase Chain Reaction
pers.comm.	personal communication
ppm	parts per million
PTGS	Post-Transcriptional Gene Silencing
RdDM	RNA-directed DNA Methylation
RE	Restriction Enzyme
RNA	Ribonucleic Acid
RNAi	Ribonucleic Acid interference
rpm	revolutions per minute
RPM	Reads per Million
s	silent
sec	seconds
siRNA	short interfering Ribonucleic Acid
TAE	Tris base, Acetic acid and EDTA
TE	Transposable Element
TGS	Transcriptional Gene Silencing
UV	Ultra Violet
v	variegated
WT	Wild-type
w/v	Weight to Volume ratio
YFP	Yellow Fluorescent Protein

1 General Introduction

1.1 General Overview

One of the current definitions of 'epigenetic' is the meiotically and/or mitotically heritable change in gene expression that is not attributed to a change in DNA nucleotide sequence (Wolffe and Guschin 2000). This refers to a second layer of information held by the genome 'on top' ('epi') of DNA. However, the origin of this term, as coined by C.H.Waddington, defines epigenetics as "the branch of biology which studies the causal interactions between genes and their products, which bring the phenotype into being" (Waddington 1942). Time has modified this definition to include phenomena that could not be explained by Mendelian inheritance. The current definitions take into account common mechanisms that have been found to underpin these phenomena, which include cytosine DNA methylation and the post-translational modification of histones. These define the chromatin environment and ultimately gene expression in eukaryotes.

Epigenetics is therefore involved in the basis for cell-specific transcriptomes, differential gene expression and the transcriptional regulation of foreign genomic elements in multicellular organisms. Furthermore, different epigenetic genome regulatory mechanisms between plant species are thought to produce different genome structures, hence differential size and complexity of genomes. The *Arabidopsis* genome contains approximately 27,000 genes and between 20-25 Mb of Transposable Elements (TEs), whereas the *Zea Mays* (maize) genome possesses approximately 40,000 genes and over 1800 Mb of TEs (Baucom et al. 2009; Schnable et al. 2009). Recent work suggests that TEs and regions of repetitive DNA have a role to play in the dynamics of genome structure, function and evolution of plants (Baucom et al. 2009; Fedoroff 2012). Due to the size of the maize genome, distribution of TEs and greater detection of epigenetic phenomena, such as paramutation, (see 1.2.3) maize is termed to have a 'complex' genome (over plants with smaller less complex genomes, such as *Arabidopsis*) (Fedoroff 2012). The repetitive nature of a complex genome such as maize has proved excellent for the study of paramutation, which allows for non-mendelian gene interactions. However, the strict circumstances of DNA sequence and genome structure required for genes to behave in this way is still unclear, although epigenetic marks and repetitive DNA are heavily implicated. Moreover, recent genome wide methylation studies conducted in maize and *Arabidopsis* have shown differences reflecting the abundance of TEs (Regulski et al. 2013; Cokus et al. 2008). To establish and maintain a larger, complex genome, will require epigenetic regulatory elements that differ in function compared to plants with simpler genomes, most obviously observed through the ability to accumulate vast amounts of TEs and repeat elements.

In addition to maize possessing a complex genome it is also an extremely important to the global food economy, grown primarily for livestock feed, bioethanol and production of corn starch, and

as the third most planted crop in the world, (Klopfenstein et al. 2012). Therefore, understanding the mechanisms controlling gene regulation is essential for the ability to molecularly manipulate crops for agricultrual purposes.

The introduction to this project serves as an explanation of the modifications of DNA and chromatin structure that facilitate gene regulation, comparing plants and mammals and issues regarding complex plant genomes. This thesis focuses on how gene silencing is regulated by epigenetic mechanisms utilising transgenic reporters, the study of transgene paramutation-like events and the potential for epigenetic communication during reproduction maize.

1.2 DNA Methylation

DNA methylation is the covalent modification of cytosine with a methyl group at the 5' position, forming 5-methylcytosine (5mc), occurring in a symmetrical CG and CHG form and asymmetrical CHH (H = A, C or T) form (Urnov and Wolffe 2001; Wang et al. 2009a). Its presence is not ubiquitous across all eukaryotes, however, if an organism normally has it, interference resulting in DNA methylation inhibition is lethal in mammals and causes defects in plants. This is observed when introducing a homozygous knockout mutation into the murine methyltransferase in mouse embryos, greatly reducing genomic methylation levels, resulting in abortion (Li et al. 1992). Therefore DNA methylation is essential for the development and survival of these organisms.

1.2.1 DNA methylation patterns in the genome

Both plants and animals possess methylation in the genome, however, the patterns they display show differences. In plants, DNA methylation was initially thought to be primarily found in a symmetrical CG and CHG contexts (Gruenbaum et al. 1981). These results, gathered from the use of methylation-sensitive RE digests, did not show the whole picture, and the development of sequencing technology and bisulfite analysis (Frommer et al. 1992) has since shown that methylation is wide-spread in the genome, across both symmetric and asymmetric contexts. The model plant species, *Arabidopsis*, possesses methylation levels of 24% CG, 6.7% CHG and 1.7% CHH demonstrating that the symmetric CG form is the most abundant (Cokus et al. 2008). This is in contrast to mammals where methylation is only found in the symmetric CG context, with ~70-80% methylated in the genome (however, non-CG methylation has been seen in embryonic stem cells) (Ehrlich et al. 1982; Law and Jacobsen 2010). The distribution of methylation also differs, mammals possess regions of demethylated DNA, known as CpG islands, regions >200 bp comprising of an above average amount of CG residues (Ohlsson and Kanduri 2002). Whereas, genome-wide

analyses of methylation in plants finds that methylation is concentrated in the centromeres and repetitive regions (Zilberman et al. 2007; Zhang et al. 2006; Cokus et al. 2008). But small amounts are also observed in gene coding regions in plants, with a deficiency observed within gene promoter regions (Zhang et al. 2006). However, discrepancies have appeared between studies conducted using different methods to study plant epigenomes. This is evidenced by the presence of methylation within transcribed genic regions, with work stating that genes remain highly expressed (Zhang et al. 2006) and other work stating that genes are only moderately expressed (Zilberman et al. 2007). The difference in methylation patterns points towards different functions between plants and animals.

Furthermore, genome-wide methylome analysis has recently been completed in maize highlighting differences with *Arabidopsis*. Methylation levels are 65% CG, 50% CHG and 5% CHG showing a greater amount of CHG methylation, in addition to a greater level of methylation in general compared to *Arabidopsis* (Regulski et al. 2013; Cokus et al. 2008). Moreover, distribution of methylation showed that intergenic DNA possessed greater amounts of methylation when compared to gene bodies, due to the greater presence of repetitive DNA within intergenic regions (Regulski et al. 2013). However, as expected the highest levels of methyation were still observed in the pericentrometic regions of the chromosomes, reflecting that high methylation density concentrates in regions of high repeat DNA and low gene number (Regulski et al. 2013; Eichten et al. 2013; Wang et al. 2009b). This distribution may therefore confer greater regulation to genes residing near to or within regions of methylation as will be discussed later.

1.2.2 Establishing and maintaining methylation

The precise mechanisms defining the establishment and maintanence of DNA methylation in eukaryotes are still unknown. However, the last 20 years have generated a firm outline of the pathways and processes involved (Law and Jacobsen 2010).

Methyltransferases catalyse the modification of cytosines in the genome. First, removal of adenosyl from ATP and transferal to methionine is catalysed by a S-Adenosylmethionine (SAM) synthetase, providing the methyl-group source (SAM), which in turn is transfered to the pyrimidine ring of cytosine residues through the action of a methyltransferase enzyme (Chiang et al. 1996). The recycling of SAM in plants is accomplished by HOMOLOGOUS GENE SILENCING 1 (HOG1) which encodes a S-Adenosyl-L-homocysteine (SAH) hydrolase (SAHH), converting SAH to homocysteine and adenosine (Baubec et al. 2010). In plants the maintenance of methylation falls to DNA METHYLTRANSFERASE 1 (MET1) (a homologue of mammalian DNMT1) (plant variant

discovered in *Arabidopsis* (Finnegan and Dennis 1993) and CHROMOMETHYLASE 3 (CMT3) (unique to plants) (Lindroth et al. 2001) which act on CG and CHG positions respectively. Methylation of CHH positions are maintained *de novo,* by the action of DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2) (which also possesses the ability to place methylation of any sequence context and a homologue of mammalian DNMT3) and targeted by small RNAs (Law and Jacobsen 2010; Cao and Jacobsen 2002). Additionally, secondary small RNAs generated can also operate in spreading of DNA methylation (Daxinger et al. 2008). The action of DRM2 and targeting by small RNAs is currently the only known method of establishing new methylation (*de novo*) and is discussed in section 1.4.2.

Symmetrical methylation can be maintained through nuclear division by recognition of hemimethylated daughter strands at the replication fork and acted on by methyltransferases (Finnegan et al. 1998) (Fig.1A). This semiconservative model was first proposed by Holliday and Pugh (1975) for the faithful replication of methylation, and due to its ease of maintenance, is thought to explain the abundance of symmetric methylation over asymmetric classes (Cokus et al. 2008).

Maintenance of asymmetric methylation on the other hand requires a *de novo* process as methylation will have only existed on a single strand prior to nuclear division. This is facilitated by the RNA-directed DNA Methylation (RdDM) pathway, involving the action of small RNAs targeting methyltransferases (DRM2) to genomic regions, a system unique to plants (Wassenegger et al. 1994) (Fig.1B). Establishment of methylation is also catalysed by *de novo* RdDM mechanisms in plants (all sequence contexts) and animals (only CG methylation) allowing for the formation of new sites of methylation. A recent study conducted in *Arabidopsis* revealed the single nucleotide methylomes of 86 mutants involved in regulation of DNA methylation and gene silencing (including the mechanisms of maintenance mentioned here) (Stroud et al. 2013). This work currently serves as a community source and has shown preliminary relationships between different regulatory pathways and loss of DNA methylation in the genome.

In addition to the action of methyltransferases establishing methylation there are active and passive methods involved with the removal of methylation. DNA glycosylases facilitate the active removal of methylation, as observed in the genome-wide demethylation of the *Arabidopsis* central cell, as well as demethylation of imprinted genes (Gehring et al. 2009; Hsieh et al. 2009). Passive demethylation can also occur due to inactivation of DNA methylation pathways, resulting in loss and dilution following nuclear division. Currently four 5-methylcytosine glycosylases are known in plants; REPRESSOR OF SILENCING 1 (ROS1), DEMETER (DME), DME-like 2 (DML2) and DML3 (reviewed by He et al. 2011). Active demethylation is achieved by breaking the N-

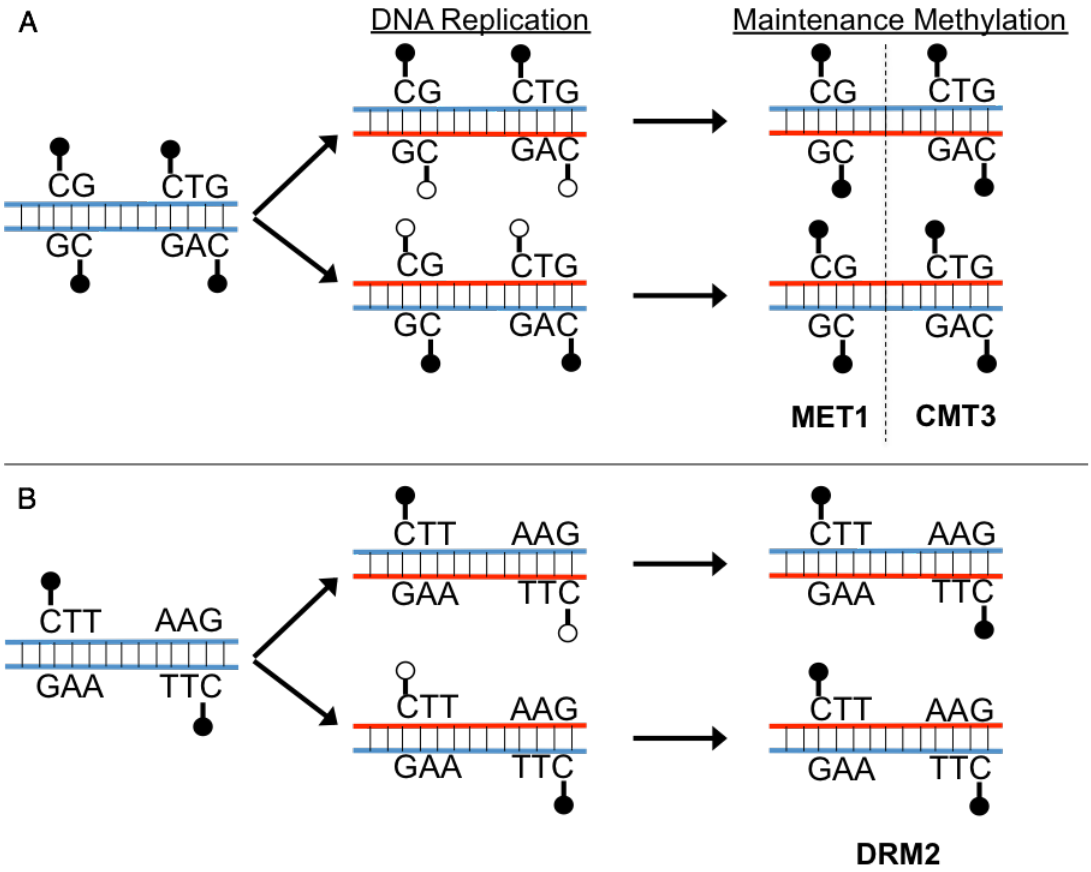


Figure 1: Diagram illustrating the maintenance of symmetric and asymmetric methylation. (A) CG and CHG (symmetric) methylation is maintained through DNA replication due to the MET1 (CG methylation) and CMT3 (CHG methylation) methyltransferases utilising hemimethylated daughter strands as a template. (B) CHH (asymmetric) methylation is maintained by *de novo* action of DRM2. Blue strands represent the template methylated DNA (filled lollipops), red strands represent newly synthesised unmethylated DNA (unfilled lollipops).

glycosidic bond, removing the base (methylated cytosine) from the DNA backbone, followed by recruitment of elements of the base-excision repair pathway for repair (Gehring et al. 2009). DME is most well known for controlling demethylation of target genes in *Arabidopsis* endosperm tissue, allowing for parent-of-origin (imprinted) expression (Huh et al. 2008). Similar enzymes also exist in mammals, functioning to erase all methylation during the production of gametes (Morgan et al. 2005; Wu and Zhang 2010). ROS1, DML2 and DML3 on the other hand are involved with the active demethylation of targets in vegetative tissues (Gong et al. 2002; Ortega-Galisteo et al. 2008). It is hypothesised that their function is to protect genes targeted for methylation, therefore ensuring their active expression.

Furthermore, the maintenance of DNA methylation, specifically CHG methylation, has been observed as the result of interaction with histone tail modifications (Histone modifications discussed in 1.3). The DNA methyltransferase CMT3 interacts with the methylated histone tails of H3K9 and H3K27 allowing for methylation of DNA (Lindroth et al. 2004). Therefore, specific histone modifications can target DNA methylation to the genome which in turn recruit further histone modifications, forming a self-reinforcing loop. The histone methyltransferase SUPPRESSOR OF VARIEGATION 3-9 HOMOLOGUE 4 (SUVH4, or KYP) provides this function, binding CHG methylated DNA and in turn allows for binding of CMT3 via a chromodomain targeted to methylated histones (Jackson et al. 2002; Lindroth et al. 2004). However, it is unknown whether protein-protein interactions occur between CMT3 and SUVH4 as observed between a similar system in mammals (Cedar and Bergman 2009). A number of other associations such as interaction between histone modifications and methyltransferases have been shown/hypothesised in plants (He et al. 2011; Saze et al. 2012). These self-reinforcing systems have raised questions over the origin of silent states, whether it lies with DNA methylation or histone modifications. Often the removal of DNA methylation is sufficient to release silencing but it remains to be shown how repressive states are established (Zhang et al. 2012).

1.2.3 The role of DNA methylation (and interaction between DNA methylation and histone modifications)

DNA methylation in the genome is generally associated with transcriptional silencing, primarily of Transposable Elements (TEs) and repeat sequences (Lister et al. 2008; Wang et al. 2009a). TEs are DNA elements of foreign origin that are regulated by epigenetics. Commonly referred to as 'jumping genes', TEs possess the ability to spontaneously 'jump' to other locations in the genome, resulting in genome mutations (Kazazian 2004). Barbara McClintock discovered TEs over

60 years ago, observing that the majority of the time TEs remained immobile (silent), revealing the first evidence of a genome regulatory mechanism (McClintock 1951). TEs and remnants of these elements comprise a large component of eukaryotic genomes (Lisch 2009), and as previously mentioned, the presence of which are speculated to have multiple roles (Fedoroff 2012).

Generally, DNA possessing high levels of methylation adopts a tightly packaged chromatin state and is referred to as heterochromatin, whereas regions with little methylation and loose chromatin is called euchromatin. It is considered to have two main roles; genome defence and the regulation of gene expression. The first role, involves protection against the transcription of foreign sequences (potentially of viral origin) and preventing mobility of TEs (Zhang et al. 2006). Especially beneficial to plants as they do not possess an antibody-based immune system (Waterhouse et al. 2001). Loss of DNA methylation through mutation of methyltransferases leads to activation of TEs and transcription of repetitive elements (Law and Jacobsen 2010).

DNA methylation is an ancient modification, the genomes answer to controlling TEs, but it would appear that in this time it has also been commandeered to function in gene regulation, its second defined role. This role includes epigenetic phenomena such as genomic imprinting, paramutation and differential gene expression in plants and involvement in early embryogenesis, stem cell differentiation, genomic imprinting and X-chromosome inactivation in mammals (reviewed by He et al. 2011). Both genome imprinting and paramutation are of particular interest to this project and are described in further detail.

1.2.3.1 Genomic imprinting Genomic imprinting describes a specific epigenetic phenomenon whereby gene expression is controlled in a parent-of-origin manner, contrary to Mendel's first law (Huh et al. 2008). This allows for maternally and paternally controlled gene expression in the next generation through transmission of different epialleles. It was originally discovered in plants in 1970, through the maternal effect of the *r1* gene, conveying colour expression in the seed aleurone (Kermicle 1970), and later in mammals (Surani et al. 1984; McGrath and Solter 1984).

In maize, imprinted genes are generally associated with reduced methylation epialleles in the central cell and sperm cells (Gutiérrez-Marcos et al. 2006). Changes in methylation create the epialleles allowing for different transmission of and subsequent expression of genes, however chromatin marks have also been implicated (Köhler and Weinhofer-Molisch 2009; Raissig et al. 2011). Studies in mammals have shown that DNA methylation essential for imprinting is removed and reestablished through each generation (reviewed by Surani 2001), whereas in plants unchanged methylation patterns are observed between generations (Kinoshita et al. 2004). Also, imprinting in flowering plants occurs most often within endosperm tissue, a product of double fertilisation that nourishes

the embryo during germination, and not within vegetative tissues (Gehring et al. 2006). The process of double fertilisation in angiosperms involves fertilisation of the haploid egg cell and diploid central cell by two haploid sperm cells, originally described in two studies (Guignard 1899; Nawaschin 1898). Both sperm cells are transported in the tricellular pollen, containing two sperm cells and a vegetative nucleus. Germination of the pollen results in formation of the pollen tube through the pistal so that the sperm is transported to the ovule. The fertilised egg cell develops into the diploid embryo and the fertilised polar nuclei of the central cell into the triploid endosperm. The 2:1 ratio of maternal to paternal genomes in the endosperm is essential for seed development, as changes to this ratio have shown abnormal development and abortion (Lin 1984). The unequal allocation of parental genomes has previously fed the theory that imprinting could be explained due the greater ratio of maternal to paternal copies.

However, numerous Maternally Expressed Genes (MEGs) and Paternally Expressed Genes (PEGs) have been identified in maize endosperm since the discovery of *r1* (Kermicle 1970), beginning with the *zein* genes and their regulators (Lund et al. 1995; Chaudhuri and Messing 1994), and more recent studies, employing high-throughput RNA sequencing technology identifying a significant number more (Waters et al. 2011; Lu et al. 2013). In addition, differentially expressed genes and DNA methylation patterns were identified in endosperm, dependent on the cross orientation of two maize ecotypes (Waters et al. 2011). Conversely, there has been very little detection of imprinted genes within embryo tissue, the first confirmed in maize was *maternally expressed in embryo 1* (*mee1*). (Jahnke and Scholten 2009). The possibility of other embryo imprinted genes is thought to be likely, Raissig et al. (2011) postulate that as DNA methylation resetting mechanisms exist in the endosperm they are also likely in embryo function, with the ability to reset a parental epiallele. Studies of maternal and paternal genome contributions to early embryo development in animals has shown that maternal gene products regulate embryo development before that of zygotic products (Tadros and Lipshitz 2009). This is also thought to occur in plants, with work carried out in *Arabidopsis* showing higher parental contributions of the maternal genome in early embryogenesis (2-4 cell embryo stage the maternal contribution is ~88%) (Autran et al. 2011). However, a further study has contested this, stating that both maternal and paternal genomes are represented equally throughout embryo development (Nodine and Bartel 2012). The authors argue that embryo RNA may have been contaminated with maternal seed coat mRNA in the previous study, attributing to their high maternal contribution. In addition, although transcripts are identified as imprinted in embryo development, they are shown to be short-lived, differing from those observed in endosperm tissues, which are more persistent (Nodine and Bartel 2012)

The purpose of endosperm imprinting (mediated by DNA methylation and/or histone modifications) is also the topic of much speculation, Gutierrez-Marcos et al. (2003) theorise a role in determining compatibility of genomes of the parents, resulting in seed abortion if unsuitable. On the other hand, Haig and Westoby (1989), define the kinship theory which proposes that imprinted genes control the allocation of nutrients to developing seeds. It has also been theorised that imprinting is essential for preventing the development of parthenogenic embryos, whereby growth can occur without fertilisation. (Huh et al. 2008). There is also no doubt that the triploid nature of the endosperm and the ratio of maternal:paternal genomes, is fundamental to its survival and development as a tissue (Gutierrez-Marcos et al. 2003).

1.2.3.2 Paramutation Paramutation refers to another epigenetic phenomenon, whereby an epigenetic state can be transferred to a homologous region elsewhere in the genome (Suter and Martin 2010). It was originally studied separately by Brink and Coe in maize, raising interest due to the non-mendelian inheritance of genetic traits (Brink 1956; Coe Jr 1959). Both studies investigated genes involved in the production of anthocyanin, the *r1* and *b1* genes. Paramutation in the *b1* locus discovered by Coe has subsequently been the most well studied, identifying two epialleles, with differing amounts of expression termed B' (silenced) and B-I (Intense). B-I can convert to B' by itself or through contact with B', therefore B' is termed paramutagenic and B-I termed paramutable. Once B-I is converted to B' it is now itself paramutagenic. The genes identified in maize all show continued transmission of a paramutagenic state following transformation and also no change in DNA sequence, indicating that the difference in expression must be epigenetic in nature.

The mechanisms for interaction between two epialleles in paramutation is currently unknown. Work conducted with *b1* has found that the presence of DNA methylation within an upstream seven tandem repeat enhancer region is a required for paramutation (Stam et al. 2002). Interestingly, this region is equally expressed in both B' and B-I epialleles, but interfering with the RdDM and the synthesis of siRNAs (discussed in 1.4.2) disables paramutation (Alleman et al. 2006). RdDM provides a method by which epigenetic states can be changed, through the action of siRNAs (Mette et al. 2000). A second theory, described as the 'pairing' model, suggests a physical association between the homologous regions, similar to that seen in *Drosophila* altering gene expression (Henikoff 1997). There is less evidence for this theory and it raises the question of how an epigenetic state could be transferred between epialleles following physical association (Suter and Martin 2010). Recent methylome analysis of two maize ecotypes has also revealed paramutation-like methylation changes in 10% of regions identified as differentially methylated, approximately

950 genes (Regulski et al. 2013). This suggests that paramutation events are far more abundant than previously thought, however, the functions of putative functions of this subset of differentially methylated regions has yet to be explored.

In addition, paramutation-like events have been witnessed affecting expression of transgenes (Khaitová et al. 2011; Meyer et al. 1993). These events are more often *trans*-silencing, originating due to sequence homology between two transgenes or multicopy insertions often involving DNA methylation (discussed in detail in 1.5.2). Although TGS often remains despite segregation of the origin silencer epiallele/homologous region (Khaitová et al. 2011; McGinnis et al. 2006), subsequent silencing has not been demonstrated. To establish true paramutation the newly epigenetically altered sequence must be paramutagenic in subsequent generations.

Nevertheless there is significant similarity between sequence homology transgene *trans*-silencing and the paramutation of endogenous genes (Matzke et al. 1996), as both systems often involve RdDM and require DNA methylation in order to repress expression. Therefore, both may trigger action usually associated with genome defence against invasive DNA (TEs) and require further investigation to dissect the mechanisms involved as well as the function in the regulation of genes (Matzke et al. 1996).

1.2.3.3 DNA methylation interactions A series of questions still exist concerning how DNA methylation in the genome is interpreted. The presence of DNA methylation is known to result in chromatin remodeling altering regulation of gene expression, defining the chromatin environment. The presence of methylation can (i) directly block the binding of transcription factors to target DNA (Prendergast and Ziff 1991), (ii) potentially alter the flexibility of DNA, changing nucleosome positioning (Segal and Widom 2009) or (iii) associate with Methyl-Binding Domain (MBD) proteins allowing for further downstream actions. It is thought that often the result of the latter two examples lead to condensation of chromatin, introducing a heterochromatic state, physically restricting access to DNA (Kass et al. 1997). However, this conclusion might prove too simplistic as there may be much about downstream mechanisms that have yet to be discovered.

MBD proteins can be divided into three classes dependent on the motifs used for recognising DNA methylation. The first class, utilises a motif named simply MBD, the first of which was identified in mice, METHYL CpG BINDING PROTEIN 2 (MeCP2), binding specifically to CpG methylated sites (Nan et al. 1993). This was followed by the discovery of interactions between MBD and histone modifying enzymes in mammals (Sarraf and Stancheva 2004). MBD domains were also found in plants showing interaction with histone modifications, currently there are 13 genes known to encode MBD motif containing proteins in *Arabidopsis* (Zemach and Grafi 2007). Different plant

MBDs identified possess different characteristics, such as universal binding at CG methylation sites by AtMBD5, AtMBD6 (associating with histone deaceytlases) and AtMBD7, or restrictive binding to CHH methylation sites, where only AtMBD5 interacts (Zemach and Grafi 2003; Saze et al. 2012). Furthermore, AtMBD9 operates as a transcriptional activator, enhancing expression of FLOWERING LOCUS C (FLC) as a result of increased histone acetylation and decreased DNA methylation (FLC expression is known to be mediated by changing histone modifications (see 1.3.3)) (Peng et al. 2006).

The second class of MBD proteins possess an SRA domain to bind methylated DNA and are associated with the direction of histone methylation, as it shares SRA domains with the SUVH family of histone methyltransferases (Johnson et al. 2007). Interestingly the plant MBD protein VARIANT IN METHYLATION 1 (VIM1) (Woo et al. 2007) and the mammalian UBIQUITIN-LIKE, CONTAINING PHD AND RING FINGER DOMAINS 1 (UHRF1) (also known as mouse Np95 and human ICBP90) appear to be involved solely in the maintenance of methylation (Woo et al. 2008; Sharif et al. 2007). Mutation of VIM (plants)/UHRF1 (mammals) results in the formation of regions of hypomethylation, not unlike the epigenetic landscape observed by the mutation of *met1* or *dnmt1* methyltransferase genes (Woo et al. 2008; Sharif et al. 2007). UHRF1 also preferentially binds hemimethylated DNA and works with mammalian methyltransferases during nuclear replication (Bostick et al. 2007). Furthermore, the plant SUVH4 (or KYP) histone methyltransferase previously mentioned, is responsible for catalysing dimethylation of H3K9, is shown to only be involved in the maintenance of non-CG methylation (Jackson et al. 2002).

The third class includes proteins utilising a zinc finger domain for recognising DNA methylation, however, examples have only been observed in mammals (Saze et al. 2012). The action of MBDs can result in the further recruitment of proteins involved in establishing different chromatin environments, such as those involved in the post-translational modification of histone proteins (reviewed in Saze et al. 2012). This opens the doors to a myriad of potential downstream interactions interpreting the presence of DNA methylation on the genome.

1.3 Histone Modification

Eukaryotic organisms package their genetic material into a complex of histone proteins forming nucleosomes, the basic unit of chromatin, the higher order structure of DNA. The nucleosome comprises of ~146 bp of DNA wrapped around an octamer of histone proteins H2A, H2B, H3 and H4 (two copies of each) (Luger et al. 1997). Each of which, possess an amino-terminal tail extending away from the core structure that can be post-translationally modified, altering protein-DNA

interaction. Modifications include methylation, acetylation, phosphorylation, propinylation, formylation, citrullation, ubiquitylation, crotonylation, sumoylation, proline isomerisation and ADP ribosylation targeting lysine and arginine residues in plants and animals (Tan et al. 2011). Adding to this complexity, individual lysine residues of histones can be, for example, mono-, di- or trmethylated. Association between specific modifications and active/silenced chromatin states have been discovered, underpinning the idea of a 'histone code' (Strahl and Allis 2000; Turner 2000), whereby combinations or sequences of histone tail modifications can produce distinct chromatin architecture associated to defined transcriptional states.

Histone modifications achieve differential behavior due to a reduction of affinity with DNA and weakened association with other histones. The tails possess a basic charge, and modifications such as acetylation or methylation are proposed to alter this charge, lessening the affinity for negatively charged DNA (Kouzarides 2007). The weakening of association between nearby histones is also thought to result in a looser chromatin structure (Norton et al. 1989). A looser, open chromatin structure is more likely to be transcriptionally active, lack DNA methylation (euchromatin), whereas tightly packaged DNA is generally silenced (heterochromatin) (Law and Jacobsen 2010). Most histone modifications are conserved across eukaryotes (Liu et al. 2010) and interact with specific structures. Chromo-like domains associate with methylation, bromo-domains with acetylation and 14-3-3 proteins with phosphorylation (reviewed by Yun et al. 2011).

1.3.1 Histone modification patterns in the genome

The distribution of modifications and their functions are not always shared between mammal and plant genomes. Classically H3K4 and H3K36 methylation is associated with active transcription and H3K9 and H3K27 methylation with repressed transcription in plants (Berger 2007). More specifically, only the trimethylated variant of H3K4 is associated with transcript abundance (H3K4me3), often found at the promoter or 5' region of highly expressed genes (Zhang et al. 2009). Whereas, both the mono- and di-methylated forms do not show a clear correlation, but are still present within genic regions (Zhang et al. 2009). H3K36 methylation is also found in regions of high transcription, specifically thought to involve itself with transcriptional elongation, linked with the working of RNA Polymerase II. This was shown due to the association with histone deacetlyation mechanisms, allowing for unraveling of DNA (Joshi and Struhl 2005).

On the other hand, H3K9me1 and H3K9me2 are associated with regions of heterochromatin, such as TEs and repetitive regions in plants (Turck et al. 2007). Interestingly the trimethylated form is often associated with euchromatin gene regions in *Arabidopsis thaliana*, the opposite of

what is observed in mammalian systems (Turck et al. 2007; Vakoc et al. 2005). The role of H3K9 methylation as a transcriptional silencer has also led to findings of direct association with the RNA-directed DNA methylation (RdDM) pathway in the maintenance of DNA methylation (Malagnac et al. 2002). Furthermore, H3K9 methylation provides a binding site for heterochromatin protein 1 (HP1) in animals (Bannister et al. 2001; Lachner et al. 2001) and the putative homologue LHP1 (like-heterochromatin protein 1) in plants (Gaudin et al. 2001). However, unlike HP1, LHP1 appears to be involved in silencing of euchromatin and not constitutive heterochromatin, through binding with H3K27me3 (Exner et al. 2009). The methylation of H3K27 also produces different chromatin states dependent on methylation. Mono- and di-methylated H3K27 is associated with heterochromatin formation, whereas trimethylated H3K27 can be found in both silenced and active gene regions in *Arabidopsis* (Mathieu et al. 2005; Zhang et al. 2007). Genome-wide microarray analysis of H3K27me3 found that its association with transcribed gene regions mostly involved transcription factors suggesting a role in the regulation of plant development (Zhang et al. 2007).

Alterations of histone modifications can also be dynamic, di- and tri-methylation of H3K9 for example are shown to be present at the beginning of transcription but quickly removed when the gene is silenced (Vakoc et al. 2005). Therefore, in addition to the histone code, sequences of changing histone modifications may be required in order to define genome operations. Recently, a study in animals reported identifying 67 new modifications, leading to a total of 130 identified to date, with a large number possibly still to be discovered (Tan et al. 2011).

1.3.2 Establishing and maintaining histone modifications

Histone modifications are catalysed by distinct enzymes, each with the ability to modify different tail locations by the different means mentioned previously. Extensive research has identified many enzymes directing modification, in addition to enzymes which actively remove modifications (for a review see Kouzarides 2007), these include classifications such as histone methyl-transferases (HMTs), histone acetylases (HACs), in addition to histone demethyl-transferase (HDMTs) and histone deacetylases (HDACs). The existence of both classes of enzyme is evidence of the dynamic nature of modifications.

The mechanisms involved in inheritance of post-translational histone modifications however, are not well understood. This involves the action of enzymes facilitating modification and also the recycling of histones during nuclear replication. A number of phenotypes dependent on the transmission of chromatin states have been observed, such as; imprinting, gene silencing, repetitive sequence/TE silencing (heterochromatin) in plants/animals and X chromosome inactivation in

mammals. Current theories state that transmission is mediated by the distribution of 'old' and 'new' histones between daughter DNA strands for a random model of chromatin or that a semi-conservative system is in place, allowing for recreation of the same histone modifications (assuming equal distribution of 'old' histones between daughter strands) (Margueron and Reinberg 2010). The maintenance of H3K27me3 has been shown to be due to association of Polycomb Repressive Complex 2 (PRC 2), which, following binding marks the site for trimethylation of the histone even on the newly formed daughter strand (Hansen et al. 2008). Another mode of transmission has also been hypothesised, whereby inheritance of H3K9 methylation is dependent on HP1 recruiting H3K9 HMTs to ensure the histone mark remains following replication (Kouzarides 2007).

1.3.3 The role of histone modifications

The main roles of histone modifications can be split into the establishment of euchromatin/ heterochromatin regions in the genome and the expression/silencing of individual genes through relaxation or condensation of chromatin. As previously mentioned methylation of H3K9 and H3K27 serve as repressive modifications for silencing of TEs and repeat regions (alongside DNA methylation) (Berger 2007). Histone modifications are also involved in the recruitment of DNA methyltransferases and histone modifying enzymes through MBD proteins associating with DNA methylation as described previously, bridging the gap between the two epigenetic modifications.

With regards to the second role, in plants one of the most well studied individual systems requiring modification of histones is the initiation of flowering. In *Arabidopsis*, FLOWERING LOCUS C (FLC) inhibits flowering and only through exposure to cold is it silenced (vernalisation) (Sheldon et al. 2008). Cold induces expression of VERNALISATION INSENSITIVE 3 (VIN3) protein, which associates with polycomb repressive complex 2 (PRC2) leading to trimethylation of H3K27, which in turn silences *FLC* (Sheldon et al. 2008). Also, histone variants show different epigenetic regulatory control, such as the presence of H2A.Z which has a negative correlation with DNA methylation (Zilberman et al. 2008; Zemach et al. 2010).

1.4 Non-coding Small RNAs and Gene Repression

RNA molecules have been found to have actions beyond that of translation of DNA into proteins and can influence the establishment of DNA methylation and histone modifications. Plants and animals use non-coding small RNAs to regulate gene expression and heterochromatin in the genome by Post-Transcriptional Gene Silencing (PTGS) and Transcriptional Gene Silencing (TGS). TGS refers to the blocking of transcription, preventing the accumulation of mRNA transcripts. PTGS

on the other hand, involves the degradation of mRNA products in the cytoplasm, preventing their translation into proteins. Also, although the same results are produced, the mechanisms are thought to be suited to different tasks, therefore, complementing each other, achieved by changes in DNA methylation and/or the chromatin environment (Fagard and Vaucheret 2000). PTGS is often reverted through meiosis, therefore not ideal for stable/heritable silencing for genome defence, which may be more suited to TGS mechanisms.

1.4.1 PTGS mediated by small RNAs

PTGS was first discovered in plants in 1998 (Waterhouse et al. 1998) and subsequent experiments identified that small RNAs 21-25 nt in size were responsible (Hamilton and Baulcombe 1999). Different classifications of small RNAs have since been designated, based on their origin of biogenesis and include micro RNAs (miRNAs) and *trans*-acting small interfering RNAs (tasiRNAs) (Fig.2). Small RNAs in the miRNA pathway are derived from dsRNA transcripts or ssRNA (transcribed from MIR genes with palindromic sequences) which folds back on itself forming a stem-loop structure (Voinnet 2009). These precursors are then cleaved by DICER-like 1 (DCL1) into their mature sizes and transported to the cytoplasm. Here, sRNAs/miRNAs associate with the Argonaute protein AGO1 and act in RNA-Induced Silencing Complexes (RISC) guiding cleavage of complimentary mRNA sequences (Voinnet 2009; Hammond et al. 2000). Contrastingly, mechanisms in animals are found to primarily inhibit translation of mRNAs through the binding of miRNAs, rather than result in mRNA degradation (Mallory and Vaucheret 2006). tasiRNAs on the other hand, are generated from transcripts which have been cleaved by the action of miRNAs, converted to dsRNA by the RNA-dependent RNA polymerase RDR6, cleaved by DCL4 and finally associate with AGO1/AGO7 in RISC complexes (Eamens et al. 2008). These small RNAs have found to be involved in targeting developmentally important transcription factors, such as those involved in auxin response (Allen et al. 2005; Howell et al. 2007) and in maize this pathway is involved in regulation of TEs, ultimately controlling transition of juvenile to adult phase in plant development (Li et al. 2010).

1.4.2 RNA-directed DNA methylation

TGS on the other hand is the result of DNA methylation and new methylation in plants is mediated by the RdDM pathway, which is directed by short-interfering RNAs (siRNA) (or rasiRNA (repeat-association siRNA)) Mette et al., 2000; Wassenegger et al., 1994; Law and Jacobsen, 2010. RdDM serves to implement *de novo* methylation in the genome in all cytosine contexts, however, targeting

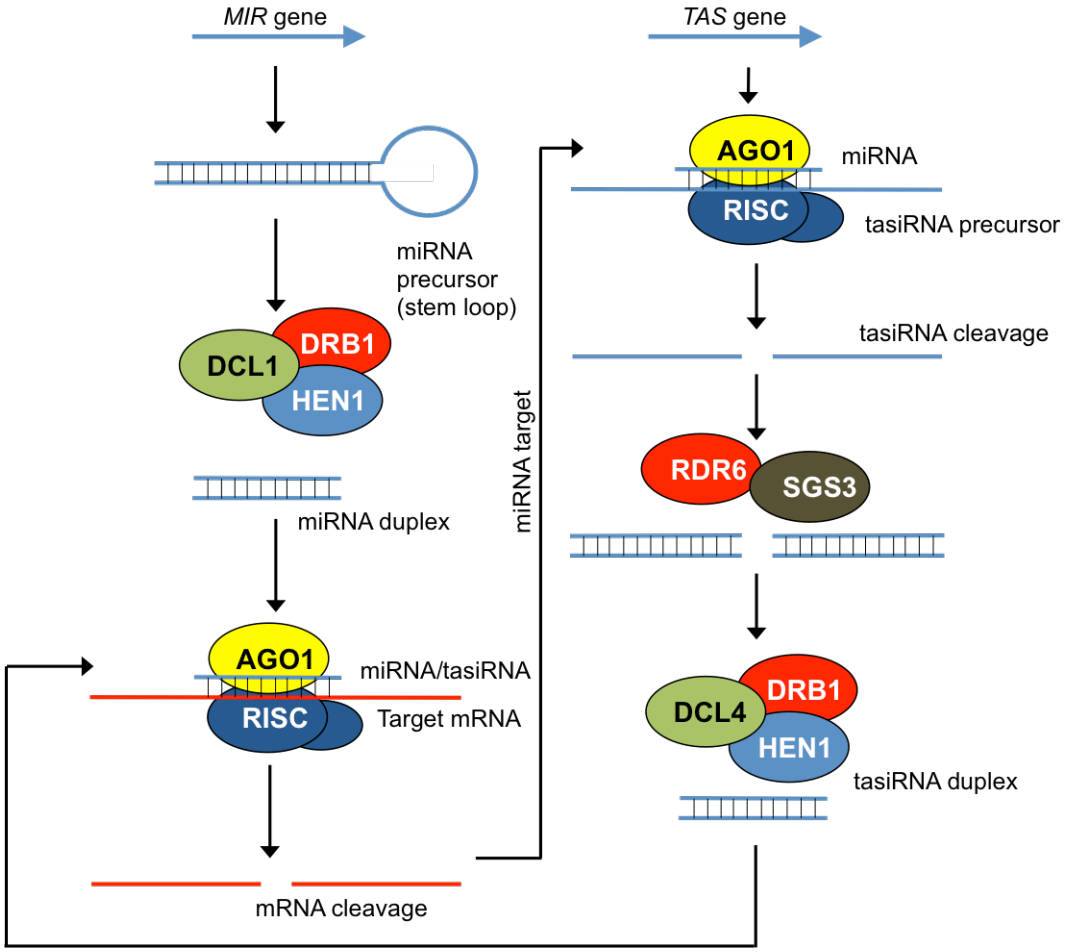


Figure 2: Diagram of Post-Transcriptional Gene Silencing (PTGS) in plants. The miRNA pathway involves the production of ssRNA transcripts from MIR genes, which fold to form stem-loop structures. These are subsequently cleaved by the action of DCL1, also involving DOUBLE-STRANDED RNA BINDING DOMAIN 1 (DRB1) and HUA ENHANCER 1 (HEN1) and finally associate with RISC complexes, at the centre of which is AGO1 resulting in cleavage of complimentary mRNA sequences. The tasiRNA pathway on right shows tasiRNA precursors originating from TAS genes targeted and cleaved by miRNAs. The ssRNA transcripts are than converted to dsRNA by RDR6 with SUPPRESSOR OF GENE SILENCING 3 (SGS3), cleaved by DCL4 and can then associate with RISC complexes and cleave mRNA targets. Adapted from Eamens et al. (2008).

directed by these siRNAs has only been observed in plants (He et al. 2011). This serves to reinforce methylation on the genome primarily for the role of TGS of TEs and repeat sequences, but also possessing roles in gene regulation. The majority of research has been conducted in *Arabidopsis*, but orthologues of the pathway have been found in other species, such as maize (reviewed by Arteaga-Vazquez and Chandler 2010). High resolution mapping of genome methylation in *Arabidopsis* has also revealed that ~37% of methylated regions are associated with high levels of siRNAs (Zhang et al. 2006). This fits with the theory that methylated regions continue to self-reinforce production of siRNAs to maintain silencing.

Dissection of the RdDM pathway has recently identified many constituents (Fig.3). In plants, ss-RNA transcripts are produced from TEs and repeat elements by the plant-specific DNA-dependent RNA polymerase IV (Pol IV) (subunits contain NUCLEAR RNA POLYMERASE D1 (NRPD1) and NRPD2), converted to dsRNA by the RNA-dependent RNA Polymerase RDR2 before being cleaved into 24 nt siRNA by DICER-like 3 (DCL3). A single strand of RNA derived from the siRNAs (the 'guide' strand) associate with Argonaute 4 (AGO4) or AGO6 or AGO9 allowing for the complimentarily targeting of DNA-dependent RNA Polymerase V (Pol V) derived transcripts (also known as intergenic noncoding (IGN) transcripts) produced from the sites of TEs (thought to be aided by RNA-DIRECTED DNA METHYLATION 1 (RDM1)) (Wierzbicki et al., 2008; He et al., 2011). Tethering between AGO4 and the IGN transcript forms the RdDM effector complex, directing the DNA methyltransferase DRM2 to place DNA methylation within the transcript region (Law and Jacobsen 2010; He et al. 2011). The chromatin remodeling protein CLASSY1 (CLSY) an SNF-2-like protein and the recent discoveries of SAWADEE HOMEODOMAIN HOMOLOG 1/DNA-BINDING TRANSCRIPTION FACTOR 1 (SHH1/DTF1) show that they are required for Pol IV recruitment and successful siRNA biogenesis (Law et al. 2013; Zhang et al. 2013). The KOW DOMAIN-CONTAINING TRANSCRIPTION FACTOR 1 (KTF1) also facilitates binding between 24 nt siRNAs and Pol V derived IGN transcripts (He et al. 2009), possibly aided by INVOLVED IN DE NOVO 2 (IDN2) (Ausin et al. 2009). The majority of other components discovered such as RNA-DIRECTED DNA METHYLATION 4 (RDM4), DEFECTIVE IN RNA-DIRECTED DNA METHYLATION 1 (DRD1) and DEFECTIVE IN MERISTEM SILENCING 3 (DMS3) are involved in aiding Pol V transcription, but are not well characterised (He et al. 2011). There are still many elements of this pathway that remain unclear, such as the role of Pol V, which is required for DNA methylation but generate transcripts without Pol IV, DCL3 and RDR2 (Wierzbicki et al. 2008). The current evidence suggests that Pol V operates independently of Pol IV derived siRNAs and that base-pairing between the two RNAs is necessary to progress the action of the pathway (Wierzbicki et al. 2008). Furthermore, recent work has shown the maintenance of

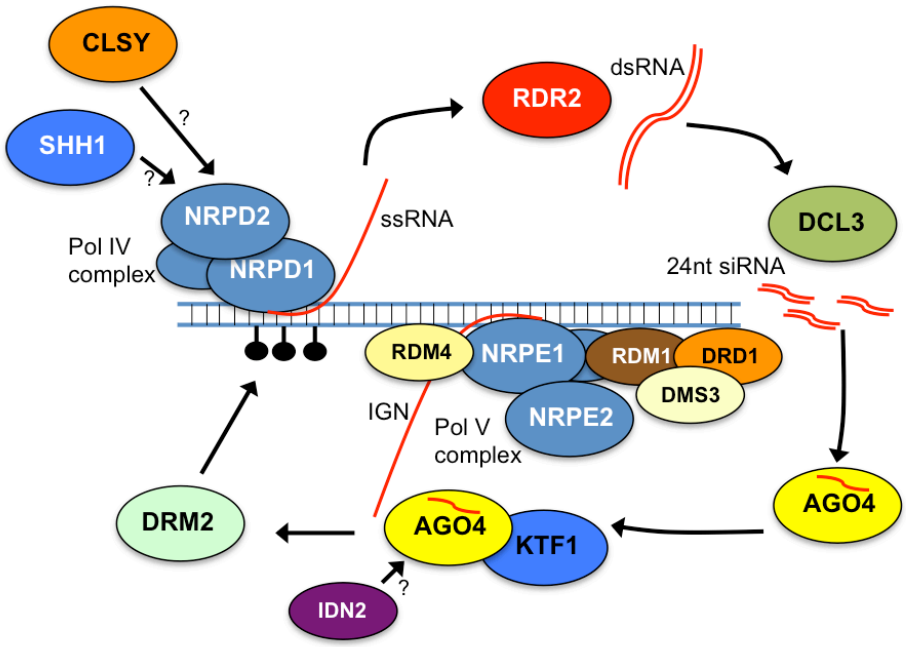


Figure 3: Diagram of the RNA-directed DNA methylation (RdDM) pathway in plants. Single stranded RNA transcripts are produced by the Pol IV complex, recruited by the chromatin remodeling protein CLSY, in addition to SHH1/DTF1. The transcript is converted to dsRNA by RDR2, cleaved into 24 nt siRNAs by DCL3 and associate with AGO4/AGO6/AGO9. A Pol V complex also generates intergenic non-coding (IGN) transcripts as a scaffold for 24 nt siRNA/AGO targeting. This requires RDM4/DMS4, DRD1, DMS3 and RDM1 for recruitment of Pol V and stable transcription. AGO4/siRNA complexes can then associate with Pol V transcripts forming an effector complex through interaction with KTF1 and IDN2 (thought to have a role in stability between 24 nt siRNA and IGN transcripts). Finally the effector complex can direct DRM2 to place DNA methylation on the genome. Adapted from He et al. (2011).

CHH methylation (however, significantly less) in Pol IV and V double mutants, suggesting other components are also involved in targeting DRM2 (Zhong et al. 2012; Wierzbicki et al. 2012).

Loss of function mutants at different positions within this pathway have shown to affect paramutation and activate silenced (trans)genes in maize, altering the expression TEs (McGinnis et al. 2006; Mosher et al. 2008; Jia et al. 2009). Interestingly, mutation of the gene encoding RDR2 results in the upregulation of DNA TEs (78%), but the downregulation of retrotransposons (68%) (Jia et al. 2009). This suggests that other mechanisms for silencing can possibly act in redundancy and target specific templates. New CHH methylation is also reinforced through attraction of further symmetrical methylation. In addition, the action of RNA Pol II non-coding RNAs derived from non-heterochromatic/siRNA producing regions have also been shown to be able to associate with AGO4 and direct RdDM machinery (Zheng et al. 2009). This system incorporates another polymerase into the RdDM pathway, with both Pol IV and V having common elements and identical or homologous with Pol II (Huang et al. 2008).

New research however, has implicated the involvement of RDR6 in placing methylation, functioning independently of Pol IV. Methylation of a single *AtCopia18A* LTR retrotransposon and a few intergenic regions were found to be targeted by this pathway (Pontier et al. 2012). Due to its independence from Pol IV, it suggests that the RDR6 RdDM pathway utilises Polymerase II (Pol II)-derived TE mRNAs that have been cleaved into siRNAs (Nuthikattu et al. 2013). The 21-22 nt small RNAs generated are usually associated with PTGS, and RDR6 has been shown to also be involved in degradation of TE mRNA transcripts (McCue et al. 2012) and could therefore provide a link between an organisms PTGS (initiation) and TGS (maintenance) abilities (Nuthikattu et al. 2013). This may provide an answer to the question of how 24 nt siRNA production is initiated for the intent of reinforcing silencing by Pol IV RdDM, which has eluded researchers in the past. This is also reinforced by evidence suggesting that 24 nt siRNAs are only transcribed from regions of DNA which are already methylated, firmly placing its role in the maintenance of heterochromatin by way of a feed-forward loop (Zheng et al. 2009; Wierzbicki et al. 2012).

1.5 Gene Silencing in Plants

Transcriptional regulation of genes is mediated by the epigenetic marks previously described, and transcription is dependent on providing a favourable chromatin environment for the binding of transcription factors with genetic regulatory elements. In plants, transgenic reporters have traditionally been used to identify the pathways and controlling elements involved in gene silencing. The mechanisms involved in gene silencing in plants are the same as those employed against DNA of foreign origin (TEs/viruses), involving both TGS and PTGS and directed by small RNAs. Similarities between TEs and certain endogenous genes/transgenes are thought to be responsible for their similar behaviour. With regards to transgenes and TEs, both are foreign to the genome, can possess very strong promoters and can be present in multiple copies (Fagard and Vaucheret 2000). However, with the differing complexity of plant genomes this response cannot be assumed to be uniform. Furthermore, research conducted in silenced transgenes has also implicated paramutation mechanisms (McGinnis et al. 2006). In maize, mutations of *mediator of paramutation 1* (*mop1*), *required to maintain repression 1* (*rmr1)* and *rmr2* all successfully reactivated a silenced *b1* genomic transgene (BTG). Interestingly, all these mutants are involved in 24 nt siRNA biogenesis (McGinnis et al. 2006). This again highlights the need for investigation into the mechanisms of paramutation.

Genetic analyses have led to the identification of two different mechanisms involved in gene silencing, (i) positional effects and (ii) homology-dependent silencing (Eichten et al. 2012; Fagard and

Vaucheret 2000). The former is proposed to cause TGS while the latter can result in TGS and PTGS as will be discussed. Following from early studies of transgene silencing, many examples of endogenous silencing events have since been shown in plants and animals, demonstrating how mechanisms involved in genome defence are utilised in gene regulatory role.

1.5.1 (i) Position-effect gene silencing

It has long been debated whether position of genes in the genome has control over transcription in plants. Positional effects were first described in *Drosophila melanogaster* and was defined as Position Effect Variegation (PEV). Tartof et al. 1984 found that the position of the *white* gene, which produces a red pigmented eye phenotype, in close proximity to regions of heterochromatin led to unstable, variegated expression (Elgin 1996) (Fig.4). This was found to be a chromosomal rearrangement (paracentric inversion), the original positioning of the white gene in proximity to euchromatin is stably transcribed, whereas, the change in position disrupts expression. This phenomenon has also been heavily studied in yeast in recent years, which has allowed for significant analysis of the epigenetic factors regulating gene expression in position-effect experiments (Chen et al. 2013). Several examples of TE/heterochromatin silencing influencing expression have also been observed in endogenous genes in plants and animals (Eichten et al. 2012; Singh et al. 2008; Matzke and Matzke 1998; Talbert and Henikoff 2006). Two examples in plants show variation of expression of FLC and BONSAI (BNS) in *Arabidopsis*. The former describes the weak action of the floral repressor *FLC* in the Landsberg *erecta* (Ler) accession, and attributes this to siRNA-mediated silencing of a TE present within an intron of *FLC* (Liu et al. 2004). The authors argue that repressive histone modifications attracted to this region lead to reduced transcriptional activity of the gene. Furthermore, recent work has found that activation of *FLC* TE by mutation of TRANSPOSABLE ELEMENT SILENCING VIA AT-HOOK (TEK), involved in maintaining heterochromatin, leads to derepression of *FLC* and a late flowering phenotype (Xu et al. 2013). The second example shows that hypermethylation and silencing of *bonsai (bns)* occurs in *decrease in DNA methylation 1 (ddm1)* mutant backgrounds, which is associated with reactivation of TEs and repetitive sequences, often leading to developmental defects (Saze and Kakutani 2007). Methylation is observed spreading from a flanking TE and is thought to be responsible for the silencing. Recent studies in maize have also looked in more detail at methylation/chromatin marks present in TEs spreading to neighbouring sequences, finding that certain families of TEs spread heterochromatin greater than others, some showing spreading of between 800-1200 bp from the TE (Eichten et al. 2012). Furthermore, genes in close proximity to TEs and repeat sequences are also thought

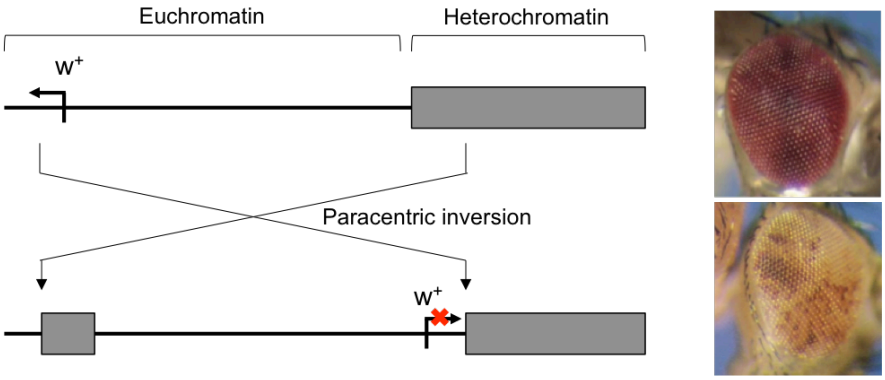


Figure 4: Schematic diagram illustrating Position Effect Variegation (PEV) in *Drosophila*. Rearrangement of position of the *white* gene (w^+) by peracentric inversion places it in close proximity to a region of heterochromatin disrupting its expression (variegated phenotype). Images show WT expression (red eye phenotype) and variegated expression (red and white expression) in *Drosophila* eyes. Adapted from Schotta et al. (2003).

to be more likely to have parental specific/imprinted expression (for a description of genomic imprinting see 1.2.3) in plants (Hsieh et al. 2009; Gehring et al. 2009). It is thought that these genes may only possess a role in the endosperm and will be silenced in vegetative tissues (Köhler and Weinhofer-Molisch 2009). Collectively, these reports indicate that the expression of genes in close proximity to TEs/repeat sequences elements can be controlled by the spread of heterochromatin. However, currently no transgene system in plants has been shown to behave in this manner, or mimic experiments conducted in *Drosophila* and yeast. A study by Singh et al. 2008 aimed to address this, utilising the regulation of Mutator (Mu) TEs in maize. Muk (Mu-killer) was used to silence MuDR elements (regulate Mu transpositions) by DNA methylation. Segregation away of Muk then usually results in stable silencing of MuDR, with Mu transposition remaining inhibited. However, a MuDR position was identified whereby induced silencing was less stable which resulted in reversion to an active hypomethylated state upon segregation of Muk (Singh et al. 2008). The failure to maintain silencing of this particular MuDR element has been attributed to position alone with DNA sequence unaffected, but the exact cause is unknown (Singh et al. 2008). They conclude that position may be important in a genomes capacity to remove methylation and other epigenetic marks. This provides the opposite role for positional effects, in the reactivation and loosening of chromatin, and not just involvement in silencing.

1.5.2 (ii) Homology-dependent gene silencing

The second mechanism known to result in gene silencing involves sequence homology. The presence of sequence homology in tandem repeat elements or at distinct loci in the genome are known to initiate silencing as part of the genomes defences against TEs (Lisch 2009; Eamens et al. 2008).

However, these mechanisms can also regulate the expression of endogenous genes. Recognition of homologous sequences within the genome has, in past studies, shown silencing following gene duplication (Rodin and Riggs 2003). A recent study of genes encoding Cysteine-Rich Peptides (CRP) in *Arabidopsis* found that some genes of the same family are methylated in a manner akin to TEs (You et al. 2012). Questions are still raised about how epigenetic machinery can distinguish between TEs and endogenous genes and the potential for an evolutionary role in gene duplication (Lisch 2009). Studies within transgenes have shown that insertion of multiple copies a transgene results in silencing (Matzke et al. 1989; Meyer et al. 1993; Mittelsten et al. 1991; Assaad et al. 1993). In addition, examples of dosage effects have been observed dependent on zygosity with silencing initiated in plants homozygous for a transgene, as opposed to hemizygous (Hart et al. 1992; Qin et al. 2003; Velten et al. 2012). The reports (mostly in *Nicotiana tabacum* (tobacco)) of these occurrences have been associated with the use of high expressing promoters, such as the Cauliflower Mosaic Virus (CaMV) 35S promoter, and have often been attributed to PTGS as a result of high mRNA production. Therefore, the doubling of mRNA caused in homozygous plants is thought to overcome a threshold leading to the silencing action (Fagard and Vaucheret 2000; Velten et al. 2012). Interestingly, in many cases homozygosity-associated silencing was only observed in a portion of total homozygous individuals, even when utilising different genes (Hart et al. 1992; Mitsuhara et al. 2002; Dehio and Schell 1994; de Borne Dorlhac et al. 1994). This suggests that the initiation of the silencing action is has a stochastic element. However, silencing as a result of homozygosity is not always stochastic, as evidenced by a recent study showing silencing of the *AtMYB90* transgene in all homozygous plants and release in all hemizygous plants (Velten et al. 2012). Furthermore, silencing would be dependent on the detection of mRNA and potential aberrant RNAs produced that can identify to the genome the location of the gene.

Repeat sequences within a sequence can also induce silencing, as shown by the direct repeats present in the promoter of the flowering gene *FLOWERING WAGENINGEN (FWA*), an imprinted gene in endosperm (Fujimoto et al. 2011). Both large and small tandem repeats present were sufficient to induce *FWA* silencing by DNA methylation in vegetative tissues, allowing for their exclusive expression in endosperm. Identification of two *Arabidopsis* accessions lacking tandem repeats, however, showed small amounts of expression in vegetative tissues (Fujimoto et al. 2008). These accessions also showed a late flowering phenotype, confirming the function of methylation of tandem repeats to stably silence *FWA* in vegetative tissues. Furthermore, the control of *b1* paramutation (for a description of paramutation see 1.2.3) in maize is dependent on seven tandem repeats located ~100 kb upstream of the *b1* gene (Arteaga-Vazquez and Chandler 2010). The repeats show DNA methylation as well as siRNAs allowing for the epigenetic transformation of

other *b1* genes. A similar set of tandem repeats, although within the coding region, has been observed in the *r1* locus, which also undergoes paramutation (Eggleston et al. 1995). However, much remains unknown about this operation, as *b1* with two tandem repeat elements does not undergo paramutation. This is thought to be due to a lack of sufficient numbers of siRNAs produced being able to act in *trans* (Arteaga-Vazquez and Chandler 2010). Both the presence of repeat sequences and paramutation have also been explored in transgenes. A study conducted by Assaad et al. 1993, successfully interfered with repeat induced TGS through targeted deletion within the repeats disrupting homology. Furthermore, paramutation-like effects, originally referred to as *trans*-silencing or *trans*-inactivation was first discovered in transgenic tobacco, when supplemented with an additional T-DNA insertion (Matzke et al. 1989). Following this addition, the original T-DNA was silenced by methylation at the promoter, which was the only region of homology between the two. Interestingly after outcrossing, segregated T-DNAs were reactivated and methylation was removed (Matzke et al. 1989). This differs from endogenous paramutation whereby the transformed loci remains paramutagenic, an attribute which currently has not been shown in plants. Other studies investigating *trans*-silencing have revealed that as little as 90 bp sequence homology can initiate silencing (Vaucheret 1993) and that variegated expression of transgenes can result from incomplete silencing by this method (Khaitová et al. 2011). *Trans*-silencing has since been shown to operate by both TGS and PTGS, directing small RNAs of multiple classes, in fact both mechanisms may operate against the same target gene (Nuthikattu et al. 2013; Velten et al. 2012; Fagard and Vaucheret 2000).

Finally the action of homologous sequences in silencing endogenous genes, through the production of dsRNA transcripts, is widespread in plants and animals. This includes the products of transcripts derived from IRs, RNA viruses and products from RNA-dependent RNA polymerases which are cleaved into small RNAs. The *Phosphoribosylanthranilate Isomerase* (*PAI*) gene in the Wasilewskija (WS) ecotype of *Arabidopsis* is one example. At this locus, an IR transcript leads to dsRNAs which are subsequently processed into small RNAs and direct local DNA methylation (*cis*) and to separate PAI loci (*trans*) (Ebbs et al. 2005). The WS-PAI locus has been employed to show the mobility of gene silencing in plant vegetative tissues (Molnar et al. 2010; Dunoyer et al. 2010a). This work showed that *PAI* small RNAs can move. This is now considered the 'gold standard' for gene silencing in plants (Velten et al. 2012). In addition, the same dsRNA result can be reached through production of sense and antisense transcripts from separate loci. Over 2,000 natural antisense genes in *Arabidopsis* are thought to operate in this way resulting in small dsRNAs that can direct PTGS against homologous mRNA (Borsani et al. 2005). Thus, gene regulation may be mediated by sequences present in the genome to prevent the accumulation of

gene products, similar to co-suppression effects found in transgenes (Napoli et al. 1990; Kanazawa et al. 2007b). Regulation can be mediated either by sequence homology between the transgene and an endogenous sequence or due to the actions of endogenous elements regulating others, such as *PAI* described earlier. In addition, TGS can be mediated through the production of dsRNAs, cleaved by DCL3 and acting in the RdDM pathway (Law and Jacobsen 2010). Furthermore, this action serves as a defence role against the introduction of viral RNA, which can be converted to dsRNA by RNA-dependent RNA polymerases and then lead to PTGS of viral gene products or viral DNA methylation (Wang et al. 2012). This may be the reason why viruses in turn target the PTGS machinery to allow for their proliferation through plant tissues (Bivalkar-Mehla et al. 2011).

1.6 Maize as a Model Organism for Gene Silencing Studies in Plants

Maize is a model monocot system in plants and compared with study of *Arabidopsis* or *Oryza sativa* (rice) has both disadvantages and advantages. The long generation times, increased size of growth facilities and increased labour amount required work against the use of maize, compared with smaller, more easily managed model organisms like *Arabidopsis* and rice. However, the genetics of maize have been well-studied over the past half century, leading to the discovery of epigenetic phenomena such as (i) genome regulation of TEs (McClintock 1951) (ii) paramutation (Brink 1956; Coe Jr 1959) and (iii) genomic imprinting (Kermicle 1970), made possible by obvious phenotypes presented by the plant. The size of seed that is produced is also an advantage, allowing for ease of tissue collection and nucleic acid extractions, and also the separation of seed component tissues. This is a process which is extremely difficult in *Arabidopsis* as seeds are very small.

Maize also has a much larger and highly variable genome, with up to 85% of its genome comprising of TEs (Schnable et al. 2009). Therefore, it is ideal for the study of endogenous gene and transgene silencing in the context of a repetitive/complex genome as previously mentioned. The size of the maize genome and the presence of TEs is thought to have a role in the dynamics of genome structure, function and evolution. A recent publication argues that greater presence of TEs in complex genomes provides greater evolvability, and that some plants have evolved their epigenetic mechanisms to accumulate TEs (Fedoroff 2012; Cokus et al. 2008). The TE insertions within intergenic regions of different maize ecotypes are known to introduce variation (Wang and Dooner 2006), conveying genome flexibility, a greater role for genome evolution than previously thought. In addition, as mentioned previously, paramutation is thought to be dependent upon the presence of repeat elements of TEs meaning that a complex genome may operates better for this form of

epiallele interaction, unlike less complex genomes, such as *Arabidopsis*.

Currently, the RdDM pathway has been shown to be responsible for co-ordinating genome structure, and although the majority of research into RdDM mechanisms has been conducted in *Arabidopsis*, novel components have also been discovered in maize. Both *required to maintain repression 1 (rmr1)* and *rmr2* shown in maize control in part the accumulation of 24 nt siRNAs, resulting in ~60% reduction in loss of function mutants and also interfere with the placement of DNA methylation (Barbour et al. 2012; Hale et al. 2009). A potential orthologue of *rmr1* in *Arabidopsis*, however, does not contain the same SNF2-like region postulated to be able to bind with chromatin marks (Law et al. 2011). Moreover, there is currently no known orthologue of *rmr2*. The effects on gene silencing that both of these components induce is essential to understanding genome regulation in maize. It is unclear how conserved components of RdDM are between differing plant species, especially as the pathway itself is still being dissected, with components such as RDM4, DRD1 and DMS3 not studied in maize.

1.7 Project Aims

The work presented in this thesis involved the molecular and genetic characterisation of a novel transgenic reporter, which exhibited a variegated expression, in an attempt to better understand epigenetic gene silencing mechanisms in plants with complex genomes. The main aims were:

- To characterise, at the molecular and genetic level, a novel transgenic reporter which exhibits a variegated phenotype in maize, an initial indication of epigenetic gene silencing.
- To determine the pathways involved in silencing of the transgene, and attempt reactivation through the use of mutant introgression.
- Potentially utilise the transgene as a tool to study other epigenetic phenomena, such as paramutation and the potential for small RNA movement between seed components, serving as an epigenetic sensor.

Maize is known to have the most complex genome currently sequenced by modern techniques, much larger than *Arabidopsis* for example. It is in this size, the result of the collection of vast amounts of TEs, that suggests the existence of different forms of genome regulation. We also know that TE heterochromatin has an effect on neighbouring DNA regions and therefore the activity of some TEs, especially those present in intergenic regions, may have a role in the regulation of gene expression.

In order to observe the mechanisms controlling gene silencing, a transgenic reporter was utilised, providing advantages over the observation of endogenous genes. Firstly, it is easy to visually observe expression through the production of a nuclear-associated YFP reporter. Fluorescent reporters are often used to detect defective transformation events, identified due to variegated or silenced transgene expression and has been well documented. Secondly, the sequence and transcriptional function of the transgene is a known factor, due to the majority of transformation events resulting in ubiquitous transgene expression. Therefore, a variegated expression hints at the existence of differing epialleles. Finally, to determine the requirements for paramutation-like events in maize, currently only transgenes containing endogenous sequences already known to be paramutagenic have worked. Therefore a transgene system with a known sequence can be used to test if *trans*-silencing can be surpassed and allow for transgenes to behave in a paramutagenic manner. This may shed light on the requirements for genes to become paramutagenic, which has so far proved elusive.

Also, this transgene has subsequently been used as a tool for the study of other epigenetic phenomena in plants such as genomic imprinting as well as the potential for small RNA movement between different components of the developing seed (described in 5.1.3 and 5.1.4).

2 Materials and Methods

2.1 General Plant Materials and Methods

2.1.1 Nomenclature

This report uses nomenclature as classified by the MaizeGBD (http://www.maizegdb.org/ maize_nomenclature.php). Gene names and mutant alleles are presented in lower-case and italics (e.g. *mop1-1*), with the nonmutant allele possessing a capitilised first letter (e.g. *Mop1*). However, if the mutant is dominant or semi-dominant then that mutant allele will be presented with a capitilised first letter (e.g. *Mop2-1*) and the wild type allele with a lower case first letter (e.g. *mop2*). Synthetic gene products are presented in all capitals (e.g. YFP).

2.1.2 Plant stocks

Plant stocks used in this project are listed in Table.1.

Table 1: Plant materials

Plant stocks	Abbrev.	Source	Accession	Reference
B73 wild type (WT)	B73	-	-	Schnable et al. 2009
Transgenics				
Nuclear YFP Reporter	NYR	Biogemma	-	-
Cell Wall YFP Reporter	CYR	Biogemma	-	-
b1 genomic transgene	BTG	Vicki Chandler (University of Arizona)	-	McGinnis et al. 2006
Epigenetic Mutants				
required to maintain repression 1-1	*rmr1-1*	Jay Hollick (UC, Berkeley)	EU154999	Hale et al. 2007
required to maintain repression 2-1	*rmr2-1*	Jay Hollick	JQ682647	Barbour et al. 2012
required to maintain repression 6-1	*rmr6-1*	Jay Hollick	NM001195895	Erhard et al. 2009
mediator of paramutation 1-1	*mop1-1*	Jane Dorweiler (Marquette University)	DQ845347	Dorweiler et al. 2000
mediator of paramutation 2-1	*Mop2-1*	Jane Dorweiler	NM001190370	Sidorenko et al. 2009
unstable factor for orange 1-1	*Ufo1-1*	Surinder Chopra (Penn state University)	un-cloned	Chopra et al. 2003
morpheus' molecule 1-like	*mom1-like*	Biogemma	AC202086	pers. comm. Biogemma
Rpd-3 like	*Rpd-3 like*	Biogemma	NM001111431	Rossi et al. 2003 De Rubertis et al. 1996

2.1.3 Plant growth conditions

Plants were grown in compost composing of 1 part Sphagnum moss peat, 1 part John Innes No.3, 1 part course grade horticultural sand and 3g/L Osmocote slow release fertiliser. Plants grown in glasshouse conditions (16-hour day/ 8-hour night photoperiod) were germinated in 9cm pots and transferred to 23cm pots after 3-4 weeks of growth and grown to maturity. Growth chambers were utilised for screening during the winter period for phenotypic analysis (16-hour day, 28°C/ 8-hour night, 22°C photoperiod, light intensity 300 μmol $m^{-2}s^{-1}$, humidity 60%, CO_2 350 ppm).

2.1.4 Maize crosses

Maize crosses were carried out by hand, individually selecting and pollinating plants. Prior to flowering the presence of ear shoots and pollen shedding were checked on a daily basis. Once ear shoots emerged they were covered with a shoot bag. It was then necessary to check daily for emergence of silks from ear shoots. The day prior to pollination, silks were cut back to 1cm in length and plants selected to serve as males have the pollen removed from the tassels. The following day a surface is held up to the male parents tassel and gently tapped to release the pollen. Once this is collected it was deposited onto the trimmed silks of the female ear shoot, which is covered again afterwards. Plants were continually watered for approximately 4 weeks, until seed development is complete, before drying the plants and harvesting ears. Seeds were stored at 5°C and 20% humidity.

2.1.5 Paper towel germination

Maize seeds were germinated directly between layers of wet paper towel inside containers to avoid evaporation and left in an a dark environment at 25°C until root emergence. Tissue could then be used in experiments or the seedlings moved to soil for further growth.

2.1.6 Seed sterilization

Seeds were first sterilized in 80% ethanol for 3 min, then 50% bleach solution for 15 min. The bleach was replaced and sterilized for a further 15 min before being rinsed 5 times in sterile Millipore water. Seeds could then be dissected or imbibed in Millipore water for 24 hours and placed on plates comprising of MS media + vitamins (4.4 g/L) (Melford), pH 5.8 (200 mM KOH), 0.9% agar (Melford) for germination.

2.1.7 Maize transformation

Embryogenic type II calli were initiated by culture of immature embryos of maize hybrid Hi-II and maintained by regular subculturing for three to six months (Armstrong n.d.). Plasmid DNA was isolated with the QIAprep spin plasmid miniprep kit (Qiagen) and coated onto tungsten (M10) particles according to Klein and Fitzpatrick-Mcelligott 1993. Callus pieces of 20 mm^2 were placed in the centre of a Petri dish with a culture medium of high osmotic pressure (0.2 M mannitol, 0.2 M sorbitol), 4 hours prior to transformation. They were co-bombarded with the e35s:ZmH2B5-YFP plasmid and plasmid pDM302 (Accession No. X17220) carrying a pat gene under the control of a rice actin promoter (Cao et al. 1992) using a particle inflow gun (Finer et al. 1992). One day after bombardment the calli were transferred to a medium of normal osmotic pressure containing 2 mg/l glufosinate–ammonium as selective agent. The calli were subcultured under selective pressure every two weeks for three months and then placed on regeneration medium to regenerate plants (Vain et al. 1993).

2.2 General Molecular Materials and Methods

2.2.1 DNA extraction

Leaf sections (approx., 2.0 cm x 0.5 cm) were collected and placed into 96-well assay blocks (Corning), with each well containing a metal ball bearing and freezing at -80°C for a minimum of 20 min. The tissue was ground by loading the assay block into a MixerMill M300 (Retsch) using a Qiagen 96-well adaptor set and grinding at 18 rps (revolutions per second) for 1 min. Sample material was spun down and 300 µl of extraction buffer (100 mM Tris-HCl, pH 8.0; 50 mM EDTA, pH 8.0; 500 mM NaCL; 10 mM β-Mercaptoethanol) and 40 µl 10% SDS were added to each well, vortexed and incubated at 65°C for 20 min. Assay blocks were incubated on ice for 10 min and then 100 µl 5M KAc was added and left to incubate on ice for a further 20 min. After a 13,000 rpm, 20 min centrifugation using a Sigma 4K15 centrifuge with Nr.09100 rotor and 09366 buckets, 125 µl of the supernatant was removed to 200 µl of isopropanol and incubated at -20°C for 30 min. Samples were centrifuged again, the isopropanol was removed and the pellet was washed in 70% ethanol. The pellet was air dried and then resuspended in 100 µl TE buffer (10 mM Tris-HCL; 1 mM EDTA; pH 8.0) supplemented with 20 µg/ml RNAseA (Invitrogen). DNA was stored at 4°C.

2.2.2 PCR reactions

PCR reactions used 5 µl DNA template (2.2.1) with a 45 µl of a mastermix comprising of, 0.5 µl each of gene specific primers (20 µM), 0.5 µl dNTPs (10 mM) (Invitrogen), 1 µl Dimethyl sulfoxide (Sigma) 1.5 µl $MgCl_2$(50 mM), 36.25 µl H_2O, 5 µl reaction buffer (10x) and 0.25 µl Taq polymerase (Invitrogen). Reactions were carried out using a PTC-225 Peltier Thermal Cycler (MJ Research) with the cycles as follows: at first a denaturation step of 95°C for 30 sec, followed by 25 cycles of denaturing at 95°C for 30 sec, anealing at a temperature appropriate for the primers for 30 sec and a 72°C extension for 1 min per kb of product. A final 72°C extension for 5 min is included at the end. Samples were then resolved by gel electrophoresis using 1xTAE agarose (Invitrogen) gels (1.5% w/v) and visualised using a UV transilluminator (Syngene G-BOX and GeneSnap software).

2.2.3 Urea DNA extraction

Tissue was collected and ground in liquid nitrogen with a mortar and pestle, 6 ml Urea EB (50 mM Tris-HCl, pH 8.0; 20 mM EDTA, pH 8.0; 300 mM NaCl; 7 M Urea; 10 mM N-Lauroylsarcosine (Sigma)) and 6 ml phenol:chloroform mix (Equal parts phenol and chloroform:isoamyl alcohol (24:1); 1 µl/ml β-Mercaptoethanol) was added to 6 g tissue and vortexed. Samples were then centrifuged at 4,000 rpm for 10 min using a Sigma 4K15 centrifuge with Nr.11150 rotor and 13350 buckets, after which the aqueous layer was removed and added to an equal amount of isopropanol . Samples were inverted to mix and centrifuged again. The isopropanol was removed and the pellet was washed in 70% ethanol. The pellet was then allowed to air dry and was resuspended in 500 µl TE buffer for 1 hour at room temperature. An equal volume of isopropanol and 1/10 volume of 3 M NaOAc, pH 5.0 was added to the samples and vortexed, followed by centrifugation at 13,000 rpm for 10 min. Isopropanol was then removed and the pellet washed in 70% ethanol, air dried and resuspended in 50 µl TE buffer (see 2.2.1). DNA stored at 4°C.

2.2.4 RNA extraction

Tissue was ground in liquid nitrogen using a mortar and pestle, 1 ml TRIzol reagent (Invitrogen) was added to 500 mg ground tissue. Samples were vortexed and incubated for 5 min at room temperature. 200 µl chloroform was added, followed by a further 2 min incubation at room temp. and then a 10 min, 10,000 rpm centrifugation at 4°C using a benchtop centrifuge. The aqueous layer was removed and an equal volume of isopropanol added. Samples were incubated overnight at 4°C. After a 10 min, 10,000 rpm centrifugation at 4°C, the RNA pellet was washed in 1 ml 70% Ethanol (made up with diethylpyrocarbonate (DEPC) treated H_2O) repeating the centrifugation

for 5 min. The pellet was then air dried on ice, and resuspended in 100 µl DEPC H_2O and stored at -80°C.

2.2.5 Hot phenol RNA/DNA extraction (endosperm and embryo tissue)

100 mg of tissue was ground in liquid nitrogen with a mortar and pestle, 500 µl of well mixed hot phenol EB (100 mM Tris-HCl, 100 mM LiCl, 10 mM EDTA, 1% SDS, pH 8.0, passed through 20 µm filter, in 1:1 ration with phenol, 80°C) was added and vortexed for 30-40 sec, before addition of 250 µl of chloroform:IAA (24:1) and repeat vortexing. Samples were then centrifuged at 13,000 rpm for 5 min using a benchtop centrifuge collecting the aqueous phase and precipitating with an equal volume of 4 M LiCl overnight at 4°C. RNA was pelleted by centrifugation at 13,000 rpm for 10 min at 4°C, collecting supernatant for gDNA extraction and resuspending RNA pellet in 250 µl DEPC H_2O. RNA was then precipitated once again with 0.1 volume of 3 M NaAcOH (pH 5.2) and 2 volumes of cold EtOH (-20°C), vortexed and incubated at -20°C for 1 hour. RNA was pelleted, through a 13,000 rpm, 10 min centrifugation at 4°C, followed by washing with 500 µl of 70% EtOH (made with DEPC H_2O) and another 10 min centrifugation. Pellet was air dried on ice and resuspended in 100 µl DEPC H_2O and stored at -80°C. gDNA extraction was facilitated by precipitation of DNA phase with an equal volume of isopropanol, gentle mixing and incubation on ice for 10 min. Samples were then pelleted by a 13,000 rpm, 20 min centrifugation, washed with 70% EtOH and centrifuged again for 5 min. Pellets were air dried and resuspended in 100 µl TE buffer. DNA was stored at 4°C

2.2.6 cDNA synthesis

Prior to cDNA synthesis a DNase digestion was performed, 10 µg of total RNA was combined with 1 µl RNaseOUT (Invitrogen), 10 µl DNaseI Buffer (10x) and 5 µl DNaseI (Roche) made up to a total volume of 100 µl using DEPC H_2O. Incubation for 15 min at 37°C followed. 1 volume of phenol:chloroform (Equal parts phenol and chloroform:isoamyl alcohol (24:1) was added and centrifuged at 4°C for 10 min, the supernatant then precipitated using 0.1 volume 3M NaAcetate (pH 5.2) and 2 volumes cold 100% Ethanol at -20°C for 1 hour. Pelleting was carried out, centrifuging at 4°C for 10 min before performing a wash in 70% Ethanol (made with DEPC H_2O), spinning and finally resuspending in 50 µl DEPC H_2O.

0.5 µg RNA and 2 µM Oligo dT18 (Invitrogen) were denatured at 65°C for 5 min and then chilled on ice for 2 min. The cDNA synthesis mix was added (100 mM DTT (Invitrogen), 5x Superscript II Buffer (Invitrogen), 10 mM dNTPs (Invitrogen), x1 RNaseOUT (Invitrogen), x1 Superscript II

RT (Invitrogen) and DEPC H_20) as per the manufacturers instruction in a total volume of 20µl and then incubated in a thermocycler at 42°C for 50 min followed by 70°C for 15 min. cDNA was stored at -20°C.

2.2.7 McrPCR

DNA was extracted (see 2.2.3) and digested by McrBC (an endonuclease which recognises and cleaves methylated DNA) as recommended by the supplier (New England Biolabs); 5 units of enzyme in a 10 µl reaction with 25 ng DNA. Following incubation at 37°C for 16 hours, the enzyme was heat inactivated at 65°C for 20 min. 5 µl of digested DNA (12.5 ng DNA) and an equal amount of undigested DNA were subjected to PCR amplification separately in 50 µl reactions. PCR products were then resolved by gel electrophoresis using 1xTAE agarose (Invitrogen) gels (1.5% w/v) running digested and undigested products in consecutive wells. Lack of product following McrBC digestion indicates methylation. Primers used in these reactions are listed in Table 3 and Table 4. All PCR reactions were performed alongside a separate reaction amplifying the unmethylated Fie2 region as a control for quality of the digested DNA.

2.3 Cloning Methods

2.3.1 Cloning of NYR-v flanking region

The genomic- NYR-v transgene junction sequence was isolated by Genome- WalkerTM technique (Siebert et al. 1995). Enzymes, adapters and adapters primers (AP1, AP2) were provided by GenomeWalker Universal KitTM (Clontech, BD Bioscience, Becton, Dickinson and Co., Europe). Specific primers were designed on the sequence of e35s:ZmH2B-YFP plasmid, using Primer 3 program (Rozen and Skaletsky 1999). Primer sequences are listed in Table.2. A first PCR amplification was performed with the adapter-specific primer AP1 and KAN.RV4 primer. The nested PCR amplification was then performed using the adapter-specific primer AP2 and KAN.RV3 primer, using as template 1 ul of 1:50 dilution of the first amplification reaction, according to the manufacturer's instructions. The PCR product was then ligated in the pGEM-T easy cloning vector (Promega), and cloned into *E. coli* DH5α competent cells (2.3.4) and sequenced (2.3.5).

2.3.2 Proof-reading PCR reactions

Products for cloning were amplified using Pyrobest DNA polymerase (Takara), a high fidelity enzyme, along with suitable primers and appropriate PCR conditions. The reaction mixture included

the following, 0.5 µl Pyrobest enzyme, 8 µl dNTPs (Pyrobest, 8 mM each), 10 µl Buffer (Pyrobest, 10x), 1 µl each of forward and reverse primers (10 µM), 1 µl template DNA and made up to a total volume of 100 µl with distilled H_20. A procedure was then carried out to incorporate an A tail overhang to the blunt ends of the amplified product, adding 1 µl dNTPs (10 mM) and 1 µl Taq polymerase (Invitrogen) and incubating at 72°C for 30 min. Amplified products were run on 1xTAE agarose gels (1.5% w/v), visualised using a UV light table, gel extracted and purified using the QIAquick gel extraction kit (Qiagen) following the manufacturers instructions.

2.3.3 pGEM-T Easy Cloning

Ligation of PCR product into the pGEM-T vector (Promega, Accession No. X65308) was carried out as follows; 3 µl PCR product, 5 µl ligation buffer (2x), 1 µl pGEM-T Easy vector and 1 µl T4 DNA ligase were mixed and kept at 4°C overnight. Transformation and growth of competent bacteria was carried out (see 2.3.4). Colonies were screened by Blue-White selection and positive transformants were picked with a sterile toothpick into 5 µl distilled H_20 before being used in a 50 µl PCR. Confirmation of the transformation utilised the plasmid specific M13 Forward and M13 Reverse primers (listed in Table.5).

2.3.4 Transformation of *E. coli* DH5α cells by heat shock

4 µl pGEM-T Easy plasmid and 40 µl od DH5α cells were mixed and incubated on ice for 30 min. A heat shock comprising of 1 min at 42°C followed by 3 min on ice was carried out before adding 1 ml SOC media (Super Optimal broth with Catabolite repression, see 7.1). Transformants were then incubated at 37°C for 1 hour of recovery before spreading onto LB (see 7.1) agar plates containing 2 µl/ml of the following, Ampicillin (Melford, 50 mg/ml), IPTG (Molekula, 100 mM) and X-gal (5-bromo-4-chloro-indolyl-β-D-galactopyranoside, Molekula, 20 mg/ml) and incubated overnight at 37°C.

2.3.5 DNA sequencing

PCR products (see 2.3.3) were cleaned using 15 µl product combined with 1 µl Shrimp Alkaline Phosphatase (USB) and 1.5 µl 1/10 dilution of Exonuclease I (Promega), which was incubated at 37°C for 30 min followed by 80°C for 10 min in a thermocycler. Sequencing reactions were then carried out using ABI Prism BigDye terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems). This involved mixing 1 µl of the appropriate primer (3.2 pmol) (listed in Table 5), 2 µl BigDye Terminator Ready Reaction Mix, PCR product and made up to 10 µl with sterile distilled

H_2O. PCR amplification was then carried out using the following conditions: 96°C for 10 sec, 50°C for 5 sec and 60°C for 4 min (25 cycles). An ABI automated sequencer (Applied Biosystems) was used for sequencing electrophoresis and subsequent DNA sequence data was analysed using SeqBuilder (DNASTAR).

2.4 PCR Methods

2.4.1 Quantitative real-time PCR

Quantitative Real-Time PCR (qRT-PCR) was carried out using MyiQ detection system (Bio-Rad) alongside MESA blue qPCR mastermix for SYBR Green assay (Eurogentec). Primers were designed for regions of interest by entering target DNA sequence into Primer3Plus software (http://primer3plus.com/), specifying a PCR amplicon between 80-130 bp, a consistent T_m between primer pairs (60-70°C), primer size between 22-28 nt and a guanine-cytosine content of between 20 to 80%. Reactions were conducted using 7.5 µl 2x SYBR green mastermix (Eurogentec), 0.15 µl (20 µM) gene-specific primers and 0.5 µl of a 1:3 dilution of cDNA (see 2.2.6) in a total volume of 15 µl. In order to reduce pipetting error a mastermix was prepared for triplicate reactions and then aliquoted as the final reaction mixture. Separate reaction mixtures were prepared for gene(s) of interest as well as the constitutive housekeeping control gene, *gapdh*. The following PCR cycles were used, 95°C for 5 min, 45 cycles of 95°C for 15 sec and 64°C for 1 min using reaction primers lised in Table 6. Data were analysed using iQ5 optical system software (BioRad) and Excel software (Microsoft). C_T (threshold cycle) values were normalised by subtracting the reference housekeeping gene C_T value (ΔC_T). The fold change in expression was established, calculating $\Delta\Delta C_T$(subtracting $\Delta C_T$of condition of interest from ΔC_Tcontrol) (Caldana et al. 2007; Livak and Schmittgen 2001). Values were then converted to $\log_2$ scale and plotted.

2.4.2 Genotyping

gDNA was extracted (see 2.2.3) and PCR was performed using primers (listed on Table.7 and Table.8) to genotype transgenes/mutants. PCR amplicons were resolved on 1xTAE agarose (Invitrogen) gels (1.5% w/v). *Mop2-1*, *rmr1-1* and *rmr6-1* mutants were digested following PCR amplification. PCR amplicons were digested in a total volume of 30 µl, incorporating 0.3 µl RE (*Mop2-1* - BsrBI (NEB), *rmr1-1* - PVUII (NEB) and *rmr6-1* - BstAPI (NEB)), 3 µl RE specific reaction buffer (NEB) and 0.2 µl BSA (20 mg/ml) (NEB) if appropriate. Reaction mixtures were incubated at the manufacturers stated operating temperatures for 2 hours and resolved on 1xTAE agarose (Invitrogen) gels (2% w/v) (Fig.5).

In addition, the presence of the *Ufo1-1* and *rmr2-1* mutants was determined visually by anthocyanin production throughout the plant. Crossing of two pigmented plants from a segregating population was conducted to produce homozygous mutant lines.

2.5 Bisulfite Genomic Sequencing

2.5.1 Bisulfite conversion of unmethylated cytosines

4 µg of DNA (extraction see 2.2.3) in 100 µl of water was denatured with 0.3 M Sodium Hydroxide at 37°C for 20 min. A super-saturated sodium bisulfite solution (pH 5.0) was prepared, gently mixing 10.8 g Sodium metabisulfite (Sigma), 1 ml of a 200 mM solution of Hydroquinone (Sigma) and 0.8 ml of 10 M Sodium Hydroxide. 1.2 ml of clear sodium bisulfite supernatant was added to the denatured DNA, gently inverted, overlaid with mineral oil, fastened with parafilm and incubated at 55°C for 16 hours. Converted DNA was cleaned using DNeasy mini-spin columns (Qiagen) – split between 2 columns, 500 µl each) following the manufacturers instructions and eluted into a total volume of 100 µl. Recovered DNA was then desulfonated by adding NaOH to a concentration of 0.3 M and incubated at 37°C for 20 min followed by the addition of ammonium acetate (pH 7.0) to a concentration of 0.3 M and mixing. DNA was precipitated by centrifugation at 13,000rpm using a benchtop centrifuge for 20 min at 4°C, then washed with 1 ml of 70% ethanol and centrifuged again for 1 min. The pellet was resuspended in 50 µl Millipore H_20 and aliquots of 3 µl were individually stored at -80°C. (modified protocol from Susan et al. 1994)

2.5.2 PCR amplification and sequencing of bisulfite modified DNA

Bisulfite treated gDNA (see 2.5.1) was PCR amplified using specific primers for the area of interest. In most cases this amplification required nested PCR to be performed using primers that had been designed within the first set to produce clear banding following gel electrophoresis, initial PCR and second nested PCR primers are displayed in Table 9. Primers were designed by selecting regions without cytosines or allocating a Y (C or T) for forward primers and a R (A or G) for reverse primers when cytosines are present or replacing cytosines with thymines in known areas of demethylation. The PCR band was subsequently resolved on a 1xTAE agarose (Invitrogen) gel (1.5% w/v), excised using QIAquick gel extraction kit (Qiagen) following the manufacturers instructions, cloned and sequenced (see 2.3.3, 2.3.4 and 2.3.5). A number of individual clones (specified in results for each experiment carried out) were sequenced to represent the average methylation state of individual cytosine positions for a single amplified product.

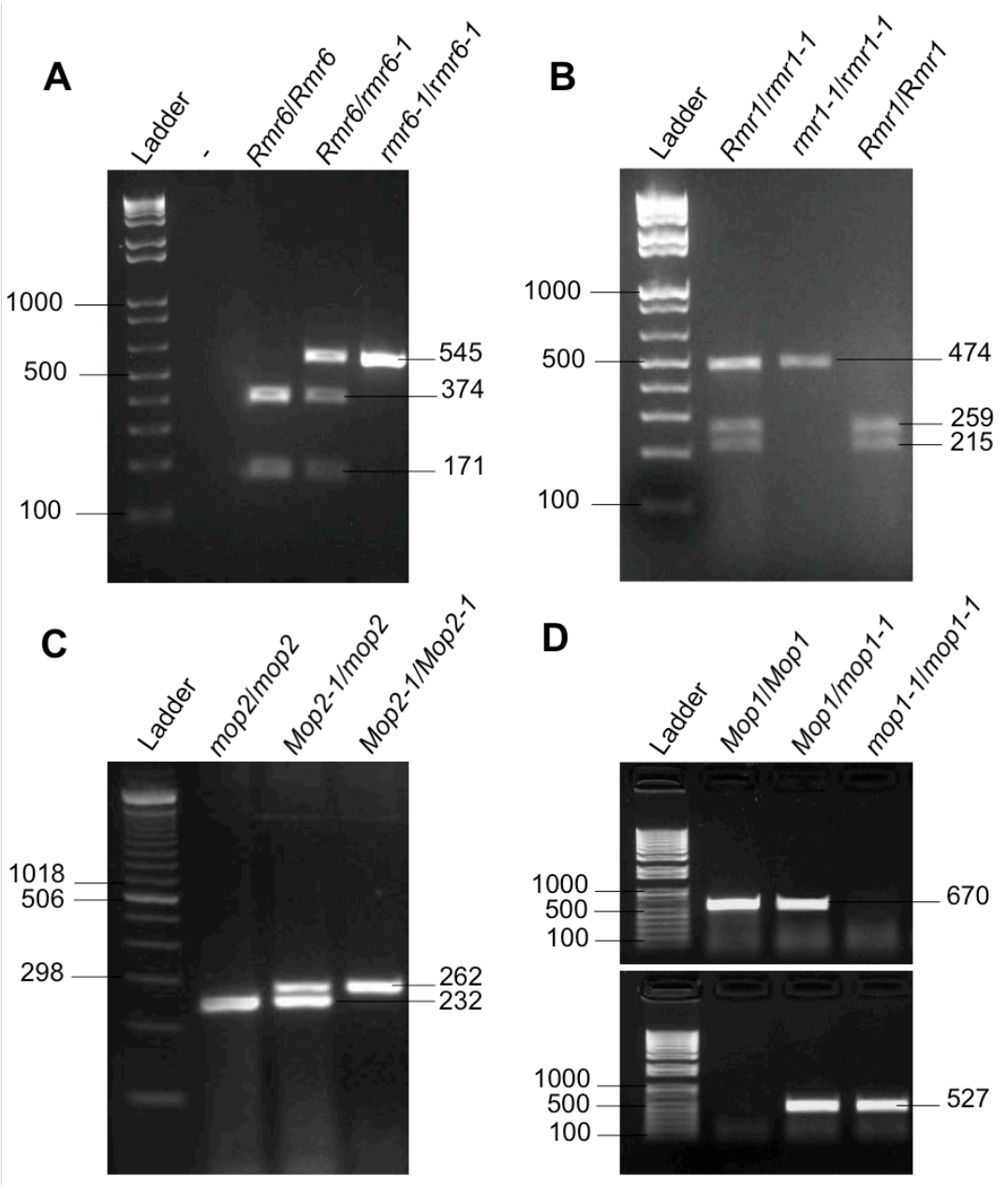


Figure 5: Genotyping maize epigenetic mutants. PCR amplicons and RE digestion revealing zygosity of (A) *rmr6-1* (digested with BstAPI RE), (B) *rmr1-1* (digested with PVUII RE), (C) *Mop2-1* (digested with BsrBI RE) and (D) *mop1-1* mutants.

2.5.3 Analysis of sequencing data

Initially, the quality of sequenced clones was assessed, disregarding sequences of insufficient length and sequence clarity. Alignment of clones was then achieved through a ClustalW alignment using MegAlign (DNASTAR) with a relevant reference sequence. Methylation at individual cytosine positions was then graphically displayed as open/filled circles across the tested region and also as a binary form for further analysis using Excel (Microsoft). Percent methylation of individual cytosine positions was calculated by dividing total number of clones with the residue methylated by the total number of clones, this was also then used to determine average methylation of the whole region and average methylation of cytosines of differing sequence context. Statistical analysis was carried out in the form of a Wilcoxon two-sample rank test (see 2.11.1).

2.6 Southern Blotting Methods

2.6.1 Southern blotting

DNA was extracted as described previously (see 2.2.3) and then 10 µg was digested with 4 µl DraI, 4 µl reaction buffer (10x), 0.4 µl BSA (20 mg/ml) (NEB) and 4 µl Spermidine (10 mM) (Sigma) in a total volume of 40 µl for 4 hours at 37°C. Digested DNA was fractionated by electrophoresis on a 1% agarose gel, covered with 0.4 M NaOH, and gently shaken for 10 min before blotting onto Hybond-N nylon membrane (Amersham) overnight. The filter was then treated with neutralisation buffer (3 M NaCl, 1.5 M Tris-HCl, pH 7.5), gently shaking for 20 min and then cross-linked using a Stratalinker UV crosslinker (Stratagene). Gel was stained with EtBr and imaged using a UV transilluminator, checking that the DNA had been transferred.

2.6.2 Hybridisation

The membrane was prehybridised for 1 hour at 65°C with prehybridisation buffer (3x SSC (20x Saline-Sodium Citrate; 3 M NaCl, 0.3 M sodium citrate (Sigma), pH 7.0), 1% SDS, 0.1% sodium pyrophosphate (Sigma), 200 µg/ml sheared salmon sperm DNA (Invitrogen) and 5x Denhardt's solution (100x Denhardt's; 10 g ficoll (GE Healthcare), 10 g polyvinylpyrrolidone (Sigma), 10 g BSA made up to a total volume of 500 ml with H_20)), followed by treatment with hybridisation buffer overnight at 65°C in hybridisation buffer (3x SSC, 1% SDS, 0.1% sodium pyrophosphate, 200 µg/ml sheared salmon sperm DNA, 5x Denhardt's solution and 10% dextran sulphate (Sigma)) and the addition of the radiolabelled probe (see 2.6.3). Membranes were then washed twice at 65°C for 30 min in 1x Wash buffer (1x SSC, 1% SDS and 0.1% sodium pyrophosphate), followed by a

further wash in 0.1x Wash buffer (0.1x SSC, 1% SDS, 10% sodium pyrophosphate), air drying, exposure to a phosphoimager cassette and analysis using a S1 Phosphoimager and associated software (Molecular Dynamics).

2.6.3 DNA probe radiolabelling

Probes were prepared from PCR products, shown in Table 10. This DNA template was denatured by boiling for 3 min and snapping to ice for 2 min before incorporation into the radiolabelling reaction mixture which included; 25 ng DNA template, 2 µl BSA (20 mg/ml), 2 µl dNTPs (0.5 µM, dCTP absent), 5 µl ^{32}P dCTP (PerkinElmer), 5x labeling buffer (Promega) and 1 µl DNA Polymerase I Large (Klenow) enzyme (Promega) in a total volume of 50 µl. Reaction mixture was incubated at 37°C for 1.5 hours before passing through a ProbeQuant G-50 micro column (GE Healthcare, following manufacturers instructions), removing any unincorporated ^{32}P dCTP. Probe was denatured at 100°C for 5 min before adding to hybridisation buffer and membrane (see 2.6.2).

2.7 Microscopy Methods

2.7.1 Laser Scanning Confocal Microscopy of YFP

Using Carl Zeiss LSM 710 confocal laser scanning microscope YFP expression was detected using 514nm excitation and emission collection between 535-565nm. Sample tissue was mounted on microscope slides with water and covered by a glass coverslip.

2.7.2 Epifluorescence microscopy of immature maize seeds

Epifluoresence imaging was carried out using a Carl Zeiss Discovery V12 microscope in conjunction with an AxioCam HRc camera (and associated software). Immature maize seeds were harvested at 8 and 12 DAP and sliced with a razor blade across the base allowing for imaging of the endosperm and embryo. YFP expression in the seeds was scored, falling into three classes; high, low and no YFP expression based on a positive YFP expressing and negative YFP seed benchmarks.

2.7.3 Laser scanning confocal microscopy of NYR and CYR expression

Root samples were analysed as in 2.7.1, imaging four locations from roots of four NYR-v/-; CYR-v/- plants. Active expressing nuclei and cell wall YFP signals were counted from each location, noting when a cell possessed both. The average count of cells with both expressing was then displayed as a percentage of total expressing cells (NYR or CYR), for each plant.

2.8 Illumina Sequencing of NYR-v Small RNAs Methods

2.8.1 Small RNA enrichment

100 µg total RNA (see 2.2.4) was used in conjunction with mirVana miRNA Isolation Kit (Invitrogen) following the manufacturers instructions enriching for RNA species <200 nt in size.

2.8.2 Small RNA library production

Small RNAs were enriched for (see 2.8.1) from 20-30 µg of total RNA extracted from leaf material of a NYR-v transgenic plant. 5' adaptor ligation was performed with DNA/RNA hybrid adaptors using T4 RNA ligase (New England Biolabs) prior to size-selection, isolation and purification from 15% polyacrylamide/Bis-acrylamide (Amresco) 3 M urea denaturing gel electrophoresis. Barcoded 3' adaptor ligation was then performed using RNA/DNA hybrid adaptors and T4 RNA ligase prior to a second round of size-selection, isolation and purification (10% polyacrylamide denaturing gel). The resulting single stranded hybrid assemblies were reverse transcribed using reverse transcriptase (Invitrogen) and a 3' adaptor specific primer. Adaptor-specific primers and the single-stranded cDNA were then used to amplify the final libraries equally with 19 cycles on a thermocycler prior to a final round of size-selection, isolation and purification (10% polyacrylamide non-denaturing gel). The Final library isolations were assessed for purity and concentration using the DNA 1000 Agilent 2100 bioanalyser and sequenced (36 bp single ended run) using an Illumina GAIIx Genome Analyser operated at the University of Warwick by in house staff.

2.8.3 Data analysis of siRNA aligning to the NYR-v transgene

Data collected from processed small RNA libraries (see 2.8.2) was output in a FASTQ format and was initially screened for barcoded sequences, using a Perl script to separate multiplexed samples. The frequency of unique reads (18-27 nt) were recorded and aligned using the short read aligner - Bowtie (Langmead et al. 2009). The small RNA library was aligned to to the NYR-v transgene specific sequence, resulting in an output showing location of small RNA alignment which was translated to normalised frequency (RPM) of small RNAs. The output was exported to Excel (Microsoft) allowing for frequency and size classes of small RNAs in different regions of the transgene to be determined.

2.9 Tissue Grafting Methods

2.9.1 Maize endosperm and embryo grafting

Following appropriate crosses, immature 6 DAP and 10 DAP seeds were sterilised (see 2.1.6) and embryo and endosperm material was separated. Thin 'slices' of endosperm tissue were cut and placed onto MS agar (0.9%) plates followed by embryos placed on top of the endosperm slice, the meristem area in direct contact. Plates were sealed and grown at 25°C in the dark. Following germination, seedlings were moved to 20 cm high containers containing the same MS agar media and grown until they reached the height of the container. Seedlings were then transplanted to soil and grown in the glasshouse until maturity, each initially individually wrapped with a plastic bag to maintain high humidity.

2.10 DNA Methylation Inhibition Methods

2.10.1 Treatment of maize roots with DNA methylation inhibitors

Maize seeds were germinated on paper (see 2.1.5) and once the tap root had reached a length of 1 cm seeds were placed into centrifuge tubes containing 1 ml of H_20, Zebularine (80μM) (Sigma) or DHPA (dihydroxypropyladenine, 200 μM) (donated by Ales Kovarik, Academy of Sciences, Brno, Czech Republic) (Fig.6). Nuclear YFP root expression was assessed by confocal microscopy (see 2.7.1) prior to this. Tubes were wrapped in foil and angled such that root tips were contacting liquid and seeds remained dry. Growth at 25°C followed and YFP expression was assessed in the root tip at 3 and 7 days of treatment.

2.10.2 Treatment of imbibed seeds with DNA methylation inhibitors

Maize seeds were imbibed in a 50ml centrifuge tube containing, 25 ml of H_20, SMZ (Sulfamethazine, 50 μM, 100 μM) (Sigma), Zebularine (40 μM, 80 μM) (Sigma), DHPA (100 μM, 200 μM) (donated by Ales Kovarik, Academy of Sciences, Brno, Czech Republic) or 5-Aza (5-azacytidine, 5 mM, 10 mM) (MP Biomedical), rocking gently in the dark at room temperature for 16 hours. Seeds were then washed with distilled H_20 before being germinated on wet paper towel (see 2.1.5). Following germination root YFP expression was analysed by confocal microscopy.

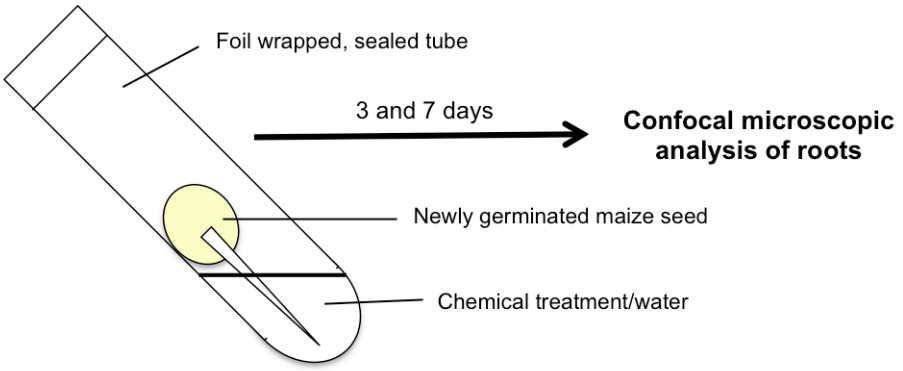


Figure 6: Schematic diagram of root treatment with DNA methylation inhibitors. The roots of newly germinated NYR-v seeds are submerged in DNA methylation inhibitor solutions/water for 3 and 7 days (tubes angled so that seed remains dry) before confocal microscopic analysis observing for NYR-v reactivation. Tubes were wrapped in foil as chemicals were known to degrade when in contact with light.

2.11 Statistical Tests

2.11.1 Wilcoxon two sample rank test

Data analysis of bisulfite data was carried out by utilising a Wilcoxon two sample rank test, a non-parametric statistical test, which, does not assume normally distributed data, using the statistical software package - R. The proportion of methylated cytosines was generated through comparison of the number of clones sequenced and for each individual cytosine position of the analysed region. Rejection of the null hypothesis occurred with a p-value of ≤ 0.05.

2.11.2 Student's t-test

Data collected from the frequency of endosperm showing active or silenced derived from maternal and paternal transmission of NYR-v were analysed using Student's t-test. This test assessed the significant difference in frequency between these classifications. All statistical calculations were conducted using Excel (Microsoft) software.

2.12 Tables of Primers

Table 2: Genome walking primers

Primer Name	Sequence (5' - 3')
KAN.RV3	GGGTGGAGAGGCTATTCGGCTATG
KAN.RV4	GCATGATTGAACAAGATGGATTGC

Table 3: McrPCR reaction primers

Gene	Primer Name	Sequence (5' - 3')	Size (bp)	Accession
UPN region	KAN.FW	AGTTCATTCAGGGCACCGGACAG	560	-
	KAN.REV	CACCGATCGCCCTTCCCAACAGTT		
e35s	35s.FW	GCGGCCGCGTTAACAAGCTTCTGCA	132	-
	35s.REV	CCACCTTCCTTTTCCACTATCTTC		
yfp	YFP.FW	CAAGGACGACGGCAACTACAAGACC	408	-
	YFP.REV	AGCTCGTCCATGCCGAGAGTGATCC		
fie1	FIE1.FW	TACTTCACAGCATGCCGCCTTCC	350	AY150645
	FIE1.REV	TTCACAGCATGCCGCCTTCCAAAGC		
fie2	FIE2.FW	GAGATTCGATTTGAAGTGTGGAAGTCC	204	AY150646
	FIE2.REV	TGATTCTCCTTGTTATGCACCGGCAG		

Table 4: McrPCR primers in analysis of NYR-v by BTG-a cross analysis

Gene	Primer Name	Sequence (5' - 3')	Size (bp)
BTG-a 35S	e35sFOR1	CTGCAGGTCCGATTGAGACTTTTCAACAA	635
	adh1REV1	CACACGGTAATAGCAGTGCTGGAC	
NYR-v e35S	e35sFOR1	CTGCAGGTCCGATTGAGACTTTTCAACAA	930
	HSP70REV1	AGTGCCCTATAACACCAACATGTGC	

Table 5: Sequencing Primers

Primer Name	Sequence (5' - 3')
M13.FW	TGTAAAACGACGGCCAGT
M13.REV	GGAAACAGCTATGACCATG

Table 6: Primers used in Real-time qPCR

Gene	**Primer Name**	**Sequence (5' - 3')**	**Size (bp)**
gapdh	GAPDH.FW	GTCACAGATGGTAGCAGGAAGGGAAG	109
	GAPDH.REV	GTGTATGCCGAGAATAAATGTGGATG	
yfp	qYFP.FW	GTGAACCGCATCGAGCTGAAGG	91
	qYFP.REV	CGTTGTGGCTGTTGTAGTTGTACTCC	
L-UPN	qKAN.FW	TCTTGTTCAATCATGCGAAACGAT	109
	qKAN.REV	CTGCAAAGTAAACTGGATGGCTTTCT	

Table 7: Primers for genotyping transgenes

Gene	**Primer Name**	**Sequence (5' - 3')**	**Size (bp)**
NYR-v	e35sFOR1	CTGCAGGTCCGATTGAGACTTTTCAACAA	930
	HSP70REV1	AGTGCCCTATAACACCAACATGTGC	
BTG-a	e35sFOR1	CTGCAGGTCCGATTGAGACTTTTCAACAA	635
	adh1REV1	CACACGGTAATAGCAGTGCTGGAC	
CYR-a	YFPFW1	GTGAACCGCATCGAGCTGAAGG	408
	YFPREV1	CGTTGTGGCTGTTGTAGTTGTACTCC	

Table 8: Primers used for genotyping epigenetic mutants

Gene	Primer Name	Sequence (5' - 3')	Size (bp)	Accession
rmr1	RMR1.FW	GCATCTTCGCAAGTTCTTCA	473	EU154999
	RMR1.REV	TCGTGGGAAGTCATCTCCTC		
rmr6	RMR6.FW	GAGGGTTTGAATCCATTGG-AATGTC	545	NM001195895
	RMR6.REV	CTTCTAGGGATGCATATTCC-AGACC		
mop2	MOP2.FW	ACAGTTGCACAGGGCTGGTT-ATTC	262	NM001190370
	MOP2.REV	AGCTGACCCATGGGCGAGGA-GGCAGTCCCGCT		
mop1	MOP1.FW	TCTCCACCGCCCACTTGAT	670	JQ248126
	MOP1.REV	ATGGCCAGCAGGGTGTCGCA-GAT		
	MOP1.FW	TCTCCACCGCCCACTTGAT	527	DQ845347
	MOP1MuTIR.REV	AGAGAAGCCAACGCCAWCGC-CTCYATTTCGTC		
mom1-like	CHR120.FW	TTCGCACCTCGGAGGAAC	$\sim$ 200	AC202086
	Mut.REV	CTTCGTCCATAATGGCAATT-ATCTC		
rpd3-like	HDA108.FW	AGACTACTACTACGGGCAAG	$\sim$ 300	NM001111431
	Mut.REV	CTTCGTCCATAATGGCAATT-ATCTC		

Table 9: Primers used following bisulfite modification

Gene	Reaction	Primer Name	Sequence (5' - 3')	Size (bp)
UPN region		delta kan FW	GTTATTTAGTTTATTTTGTA-GGGTTTTTTAATTTTATTAG-AGG	556
		Bis rv1	CCACCTTCCTTTTCCACTAT-CTTC	
(NYR-v) e35S	1a	delta kan FW2	GGATGTGCTGCAAGGCGATT-AAGTTGGG	844
		Bis rv4	CTAAACCACTCTCAGCAATC-ACCACACAA	
	1b	delta kan fw2	AGTTCATTCAGGGCACCGGA-CAG	719
		Bis rv3	TCCTCTCCAAATGAAATGAA-CTTCCTT	
(BTG-a) 35S	1a	B 35S BS FW3	GTATGTTGTGTGGAATTGTG-AGTGGATA	617
		B 35S BS REV3	CAACAAATCATAAACCAAAA-TTAATC	
	1b	B 35S BS FW2	AGTTTTTAGAGATTYGTYAA-YATGGTGGAG	526
		B 35S BS REV3	CAACAAATCATAAACCAAAA-TTAATC	
fie2		Fie2E1BS.for	AAGATTTGAGATTYGATTTG-AAGTGTG	226
		Fie2 REV1 BS	ACTTTCCCCTCCRCCTAATT-CTCCTTA	

Table 10: Primers used for southern probe analysis

Gene	**Primer Name**	**Sequence (5' - 3')**	**Size (bp)**
e35S	e35S.UN FW	GCGGCCGCGTTAACAAGCT-TCTGCA	665
	e35S.UN REV	ATCTGCTAGAGTCAGCTTG-TC	

3 Molecular Characterisation of NYR-v

3.1 Introduction

3.1.1 Chapter Aims

The aim of this chapter is to characterise the silencing action of a variegated transgenic reporter – Nuclear YFP Reporter (NYR) (Referred to as NYR-varieagated (NYR-v)), to investigate gene silencing mechanisms in maize. Maize possesses a complex genome with epigenetic mechanisms regulating gene expression that are not fully understood. Through the analysis of the silencing action on a transgenic reporter (NYR-v) it is possible to determine the mechanisms required for repression, which is the subject of the next chapter. However, initially NYR-v must be characterised, demonstrating its repressed expression at the molecular level and investigate its spatial and temporal distribution of expression. To this I set three objectives:

- Characterise the silencing of the NYR-v transgene.
- Define the molecular mechanisms responsible for NYR-v silencing.
- Define the spatio/temporal distribution of NYR-v silencing.

The following sections prior to the results describe in detail some of the tools to be used in this chapter and through the rest of the project.

3.1.2 Microscopic analysis of fluorescent reporter expression

The study of gene expression has been greatly aided by the addition of fluorescent labeling of gene products and advances in imaging technologies. These tools provide a non-destructive method in which to study temporal, spatial and strength characteristics of gene expression.

The development of expression labeling originated with conventional techniques of fluorescent staining. This utilises fluorchrome labeled chemicals or antibodies (immunofluorescence techniques) known to interact with specific targets in the cell (for a review of conventional techniques see Suzuki et al. 2007). The subsequent purification and analysis of naturally occurring bioluminescence proteins in *Aequorea victoria* (jellyfish) has allowed access to a biological molecule, which could be used for molecular labeling (Shimomura et al. 1962). This protein acquired the name Green Fluorescent Protein (GFP) and its 238 amino acid sequence was determined in 1992, facilitating its transformation into other organisms (Prasher et al. 1992). Fluorescence expression of GFP was first demonstrated showing expression in *C. elegans* (Chalfie et al. 1994) and *E. coli* (Inouye and Tsuji 1994), and a modified GFP was transformed and successfully expressed in the plant model tobacco (Chiu et al. 1996), following some difficulty. This paved the way for gene/fluorescent protein

fusions as a reporter for expression and has been widely used in the plant research community. Widespread usage has also led to the development of many other similar proteins, derived from GFP, with unique excitation and emission wavelengths, these include cyan (CFP), yellow (YFP) and blue (BFP) among many others (Tsien 1998). This has allowed for the simultaneous use of differing fluorescent protein reporters, distinguishing different transgenes or fusions, a feature that is demonstrated excellently in reporters used to create the 'Brainbow', a multi-coloured expression in mouse brains (Livet et al. 2007). Additional fluorescent reporters (non-GFP derived) have also been identified and used in biological research, such as luciferase (de Wet et al. 1985) and GFP-like proteins derived from *Anthozoa* corals (reviewed by Alieva et al. 2008).

In parallel with the development of fluorescence labeling techniques there have also been innovations in imaging. Traditional fluorescence microscopy depends on the use of a light source at a specific wavelength (dependent on the fluorophore) to excite fluorescence in a sample. The subsequent light emissions from this can then be separately detected at another wavelength. Emissions can be focused to optics for visual analysis or towards a camera for image capture (Paddock 2000). The development of Laser-Scanning Confocal Microscopy (LSCM) has provided numerous advantages over traditional fluorescent microscopes (modern equivalents pioneered by Marvin Minksy, Minksy 1988). The design of this imaging system rejects out-of-focus light in forming an image due to the placement of a pinhole in front of the photodetector, allowing for sharper, higher resolution imaging (Halbhuber and König 2003; Prasad et al. 2007). Previous fluorescence microscopy equipment often suffered from image blurring, as samples were viewed as one three-dimensional object, whereas LSCM allows for focusing on the sample as a series of planes (Halbhuber and König 2003). Furthermore, scanning of the electronic image translated from the photodetector provides quantitative values to fluorescence levels detected, the ability to digitally manipulate images, produce 3D representations of samples and simultaneously image multiple fluorophores within a single sample. This technology, coupled with pre-existing laser technologies in the excitation and detection of emissions from fluorescent molecules, has been and continues to be a powerful tool in biological research.

3.1.3 Transgene copy number analysis by Southern Hybridisation analysis

In order to utilise fluorescent transgenic reporters or any transgenes, they first have to be transformed into the organism. The insertion of multiple transgene copies or shared sequence homology between transgenes can result in attraction of DNA methylation and other epigenetic marks resulting in TGS or PTGS (reviewed in Meyer and Saedler 1996 and Fagard and Vaucheret 2000).

This has been documented in numerous examples in plants (Matzke et al. 1989; Meyer et al. 1993; Mittelsten et al. 1991; Assaad et al. 1993). Therefore, it is important to effectively assess insertion frequency during transformation. Detection of which can be carried out by restriction enzyme analysis coupled with Southern Hybridisation. The use of a unique restriction site within the transgene yields a single band in subsequent Southern blot analysis if there is a single insertion and multiple bands for higher copy numbers (either in tandem or separate locations and dependent on the location of the restriction site in neighbouring genomic regions) (Fig.7) (Velten et al. 2012; Yuan et al. 2007). Alternate methods to estimate copy number exist in the form of micoarray analysis (Lucito et al. 2000), genomic hybridisation (Larramendy et al. 1998) and Fluorescence *in situ* Hybridisation (FISH) (Kallioniemi et al. 1996). The use of quantitative real-time PCR (qRT-PCR) however, has become the prominent alternative to Southern analysis, first used by Bièche et al. 1998 in cancer studies of gene copy number and subsequently in plants, specifically maize (Ingham et al. 2001; Song et al. 2002). This technique is faster, less expensive and requires less DNA than those previously mentioned and can also be used in a high-throughput manner, (Yuan et al. 2007) using fluorescence of newly generated DNA strands as an indication of amount of origin DNA present (Yang et al. 2005). Opinion is split over whether qRT-PCR replaces Southern Hybridisation analysis in copy number estimation, or whether it can be utilised as an additional tool alongside (Yang et al. 2005; Bubner and Baldwin 2004). However, the time and quantification abilities offered by qRT-PCR methodologies are of greater advantage over Southern Hybrisisation based techniques.

3.1.4 The study of DNA methylation using bisulfite sequencing and methylation sensitive restriction enzyme analysis

Following the discovery of methylation marks on cytosine residues in the genome and their ability to affect transcription, methods were established for their detection. Traditionally, methylation sensitive restriction enzymes have been used in conjunction with Southern Hybridisation analysis. The use of HindIII, BamHI and EcoRI has been common place in this analysis, digesting DNA at restriction sites when unmethylated but blocked when methylated (Steward et al. 2000). Subsequent probing through Southern Hybridisation can then detect the presence or lack of product indicating hypomethylation and hypermethylation respectively. However, significant quantities of gDNA are required, techniques can be long-winded and analysis is restricted to sites recognized by methylation sensitive restriction enzymes (Moore 2001). An alternative method that shows the methylation state of individual cytosine residues was discovered in the form of bisulfite sequenc-

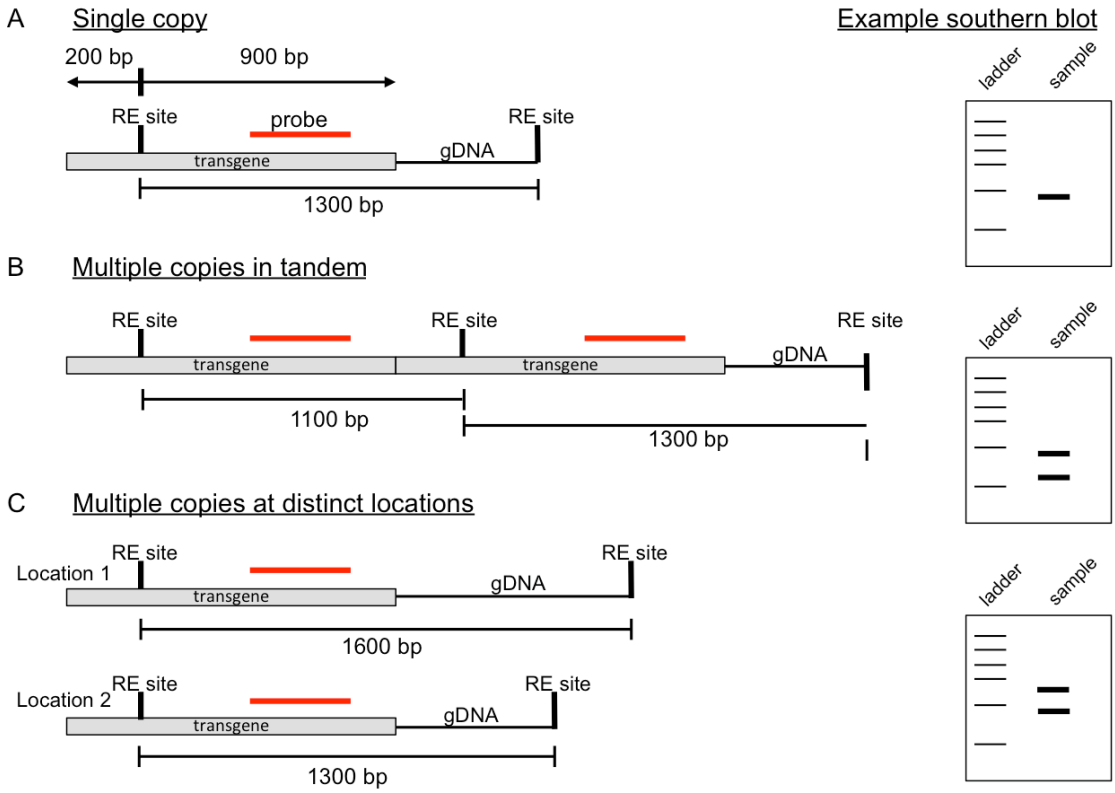


Figure 7: Schematic diagram illustrating the analysis of transgene copy number by restriction enzyme digestion coupled with Southern Hybridisation analysis. Different sizes of Southern blot hybridised products (bands) can be detected from gDNA digested at unique restriction sites within a transgene. Examples shown are (A) single copy (one band), (B) multiple copies in tandem (two bands) and (C) multiple copies at distinct locations (two bands). Example sizes of detected products are labeled below each schematic in bp. Simplification of Southern blot analysis results displayed on the right of examples.

ing (Frommer et al. 1992). Treatment of gDNA with sodium bisulfite allowed methylated and unmethylated cytosines to be distinguished, and is now the gold standard for investigating methylation (Fig.8). PCR amplification, cloning and sequencing of bisulfite converted DNA can then be aligned to a reference DNA sequence where unconverted cytosines reveal sites of methylation (Frommer et al. 1992). Due to variability in the presence of methylation, the data from multiple clones is pooled, producing an output showing proportion/likelihood of a single cytosine position possessing methylation.

In addition, other methylation analyses can produce quick results revealing the methylation state of a region. McrBC digestion followed by PCR (McrPCR) analysis is one such example (Fig.9), relying on McrBC endonuclease digestion of gDNA, recognizing only methylated cytosine residues (Sutherland et al. 1992). Subsequent PCR amplification then only occurs if the template DNA has remained undigested (therefore unmethylated). However, this analysis is not quantifiable and can only be used to detect large differences in DNA methylation. This type of analysis is useful in initial screens of regions, for which bisulfite sequencing would be too costly and impractical (Rangwala et al. 2006).

These tools enable the copy number of transgenes to be determined, their reporter expression to be observed and the reading of their methylation profiles. All of which are vital to characterise the mechanisms involved in generating NYR-v expression and observe reporter expression.

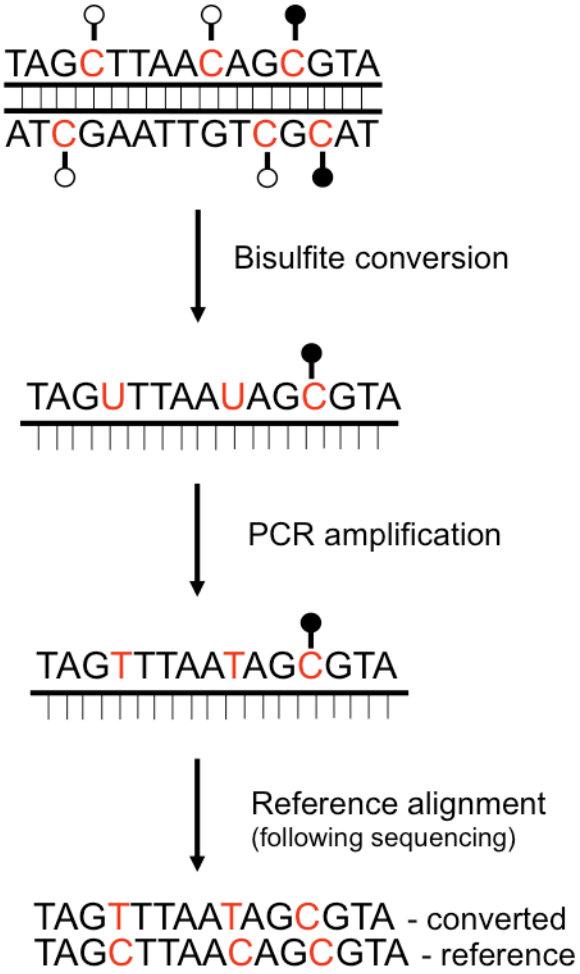


Figure 8: Schematic diagram of bisulfite modification and sequencing for DNA methylation analysis. Bisulfite treatment of gDNA converts unmethylated cytosines to uracil, while methylated cytosines remain unchanged. PCR amplification and sequencing converts uracil to thymine, and subsequent alignment to a reference sequence reveals the location of DNA methylation. Open lollipops, unmethylated; closed lollipops, methylated.

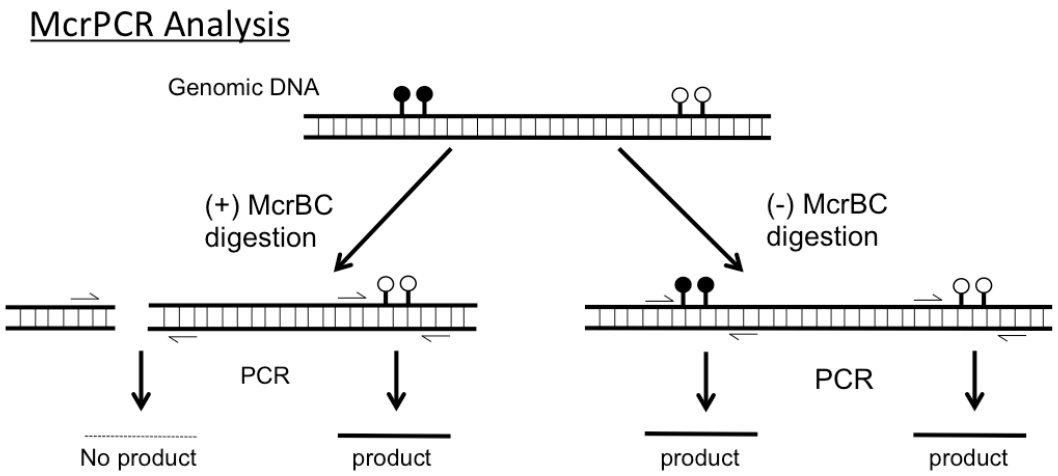


Figure 9: Schematic diagram illustrating McrPCR analysis. PCR amplification is carried out on regions of interest with an equal amount of both McrBC digested (+) and undigested (-) gDNA. A lack of PCR product will then indicate the presence of methylation, in (+) providing a product is also achieved in the undigested (-) control. Conversely, a PCR product in (+) will indicate a reduced presence of methylation. Black lollipops = methylated, Open lollipops = unmethylated.

3.2 Results

3.2.1 Generating Nuclear YFP Reporter (NYR) transgenics

The Nuclear YFP Reporter (NYR) transgene is comprised of a tandem enhanced 35S Cauliflower Mosaic Virus (e35S CaMV) promoter driving a downstream H2B histone-associated YFP reporter (H2B5:YFP) targeting the fusion protein to the nucleus (Fig.10). The e35S promoter allows for high expression of DNA of foreign origin in plant cells (Benfey and Chua 1990). In addition, in monocots the intron of Heat Shock Protein 70 (HSP70i) has been positioned immediately downstream of the promoter to improve transcription (Rochester et al. 1986). The transgene was transformed into maize B73 plants using biolistic transformation with a second construct conferring resistance to glufosinate (pDM302). Transgenic plants were screened for the presence of YFP by epifluorescent microscopy. Most transformation events showed constitutive expression of YFP, however, a single event showed a variegated expression. These transgenic plants were outcrossed over three generations, and the resulting progenies retained the variegated expression of the reporter. In this report NYR individuals with a silent or variegated phenotype are designated as NYR-variegated (NYR-v) for ease of understanding and are all derived from a single transformation event. The independent transformation events showing constitutive reporter expression are designated NYR-active (NYR-a). However, seed stocks of NYR-a were extremely limited as they were not propgated. Therefore they were only utilised in a single experiment, comparing expression levels to NYR-v plants in qRT-PCR analysis (see 3.2.5).

3.2.2 Phenotypic analysis of NYR-v in vegetative tissues

Silencing of NYR-v was observed through analysis of expression in the roots and leaf by confocal microscopy (Fig.11). Roots were selected for analysis as they are the first tissue to form from newly germinated seeds, ideal for screening. In addition, cells at the apical end of the root tip are condensed, elongating into cell files as the root grows, therefore analysis at the root tip has a considerably higher density of nuclei to observe NYR-v expression compared with larger cells in other tissues such as the leaf.

Varying levels of silencing in roots and leaves could be observed in NYR-v plants (Fig.11). Expression ranged between no visual YFP detection (rare) and higher amounts present in variegated expression. It was also evident that once a cell adopted an NYR-s (silent)/ NYR-a (active) state the phenotype was inherited mitotically to subsequent daughter cells. This is most easily observed by the cell files (lines of cells) exhibiting reporter expression within root tissue, which follows the

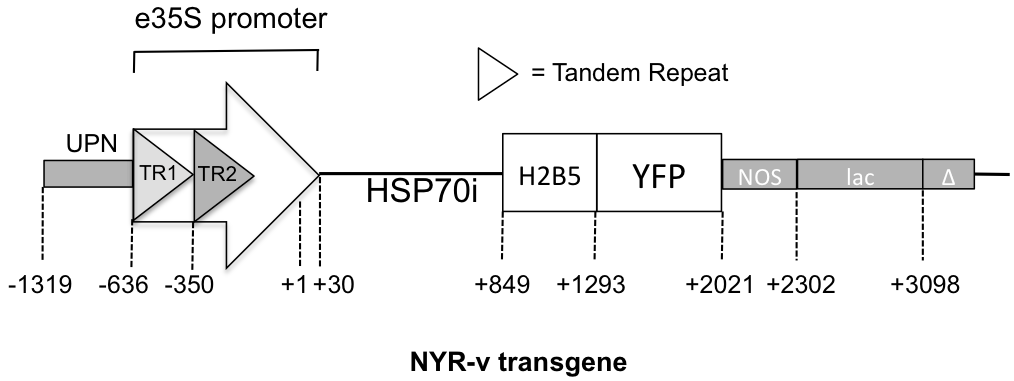


Figure 10: Schematic of NYR-v transgene. UPN, Upstream NPTII; TR, tandem repeat; HSP70, Heat Shock Protein 70; H2B5, Histone 2B5; YFP, Yellow Fluorescence Protein; e35S CaMV, enhanced 35S Cauliflower Mosaic Virus promoter; NOS, Nos terminator; Δ, delta NPTII. Positions relative to e35S transcriptional start site (+1).

growth of daughter cells.

These data show that variegation of NYR-v expression exists within both root and leaf tissues.

3.2.3 Copy number analysis of NYR-v by Southern Hybridisation

Increased copy number has often been associated with transgene silencing (reviewed by Fagard and Vaucheret 2000). Therefore, to determine the number of NYR transgene copies present, restriction enzyme digestion coupled with Southern Hybridisation analysis was utilised. The Southern Hybridisation analysis shows only a single band, indicating that NYR-v is likely a single copy insertion in the maize genome (Fig.12). In addition, it fits with the predicted size (1562 bp) of the product expected, ~1500 bp.

3.2.4 Sequence analysis of NYR-v by genome walking

Genome walking was used to establish the sequence integrity of NYR-v following transformation as well the location of the insertion site in the maize genome. Rearrangement of a transgene can occur during transformation which can impair its function or attract silencing machinery (Morino et al. 1999). Analysis revealed 1041 bp of the endogenous DNA upstream of NYR-v ((a simplified schematic diagram of NYR-v is shown in Fig.13) with the nucleotide sequence shown in Appendix Fig.71.

BLAST nucleotide analysis of endogenous DNA upstream of NYR-v utilising the maize TE database revealed the sequence as a Long-Terminal Repeat (LTR), *gypsy*-like (super-family), Flip (family) TE (RLG_flip_AC214266).

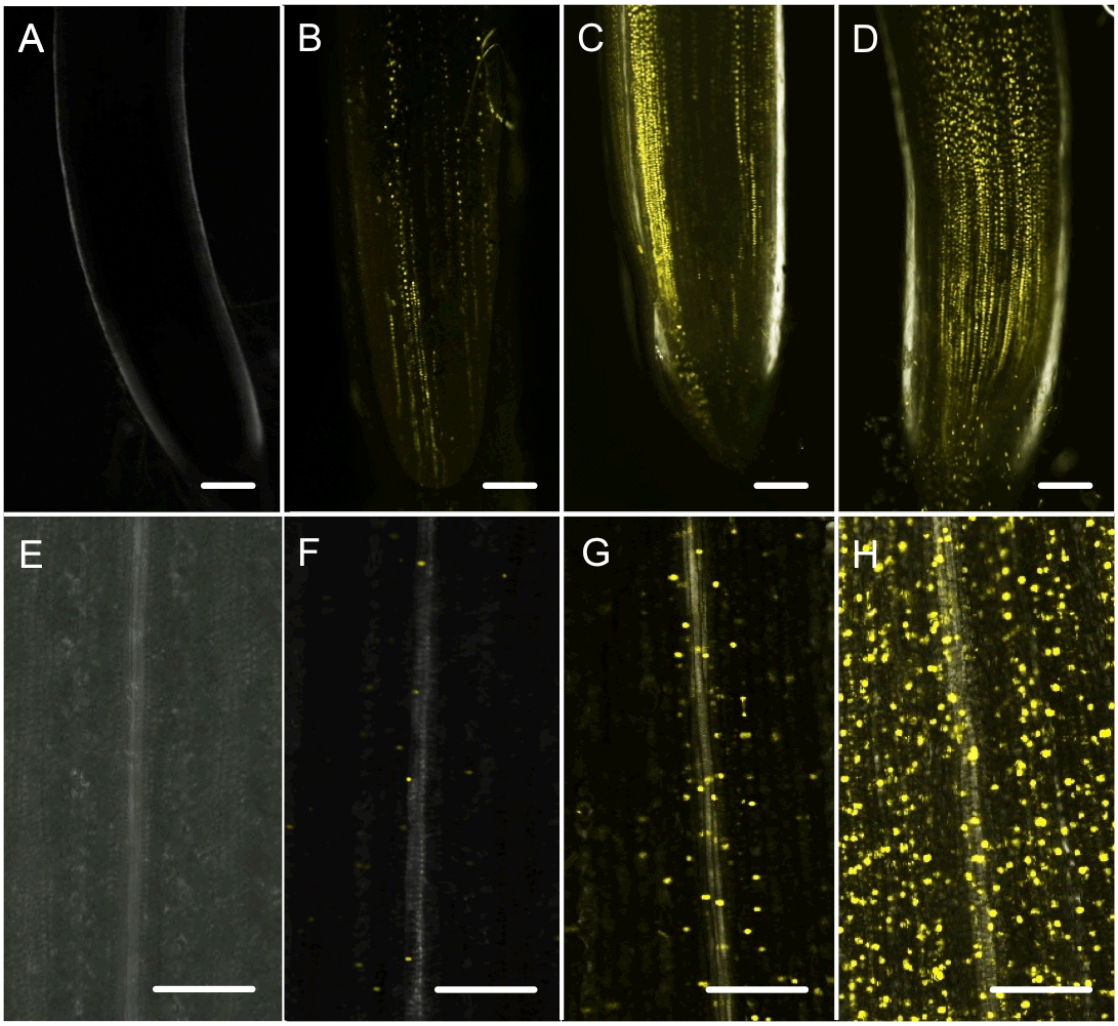


Figure 11: Confocal microscopic images of roots from NYR-v and NYR-a. Images show the NYR-v (A-C) and NYR-a (D) expression levels in root tips and corresponding leaf tissue (E-G and H). Scale bar = 200 µm.

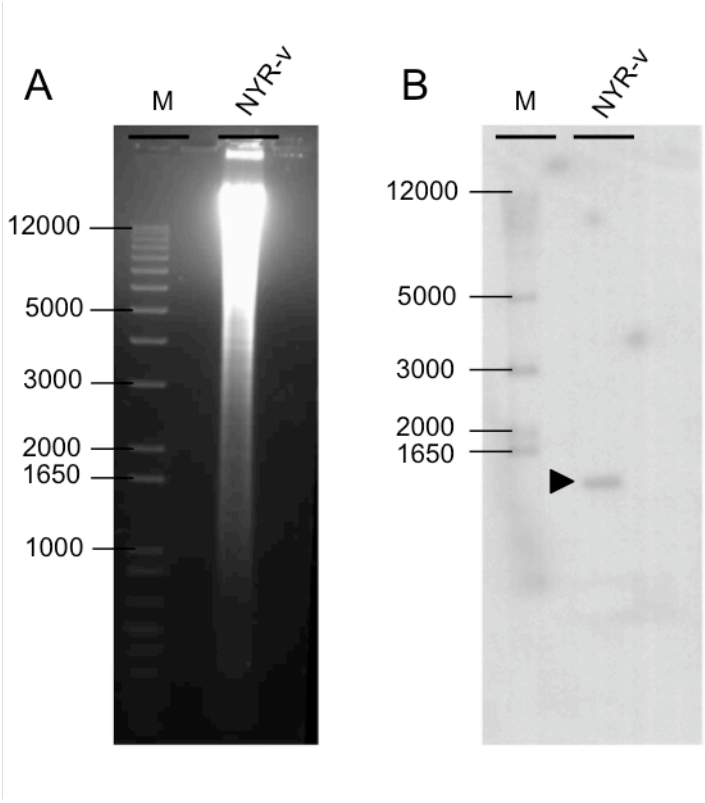


Figure 12: Copy number analysis of NYR-v by Southern Hybridisation analysis. (A) DraI digested NYR-v gDNA (B) DNA gel blot showing detection of one band. M, molecular weight DNA marker (bp).

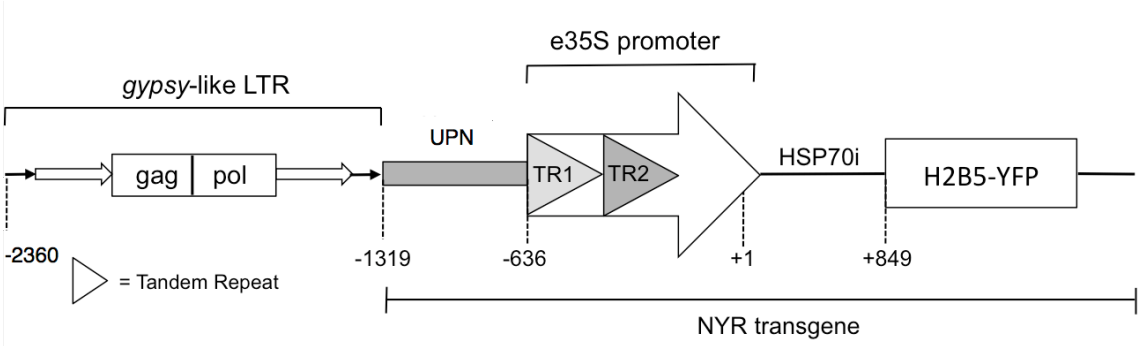


Figure 13: Schematic diagram (simplified) of NYR-v transgene with upstream LTR region. LTR, long-terminal repeat; UPN, Upstream NPTII; TR, tandem repeat; e35S CaMV, enhanced 35S Cauliflower Mosaic Virus promoter; HSP70i, Heat shock protein 70 intron; H2B5-YFP, histone H2B.5 fused with yellow fluorescent protein. Endogenous DNA upstream of NYR-v represented as *gag* and *pol* subunits of a retroelement. Positions relative to e35S transcriptional start site (+1).

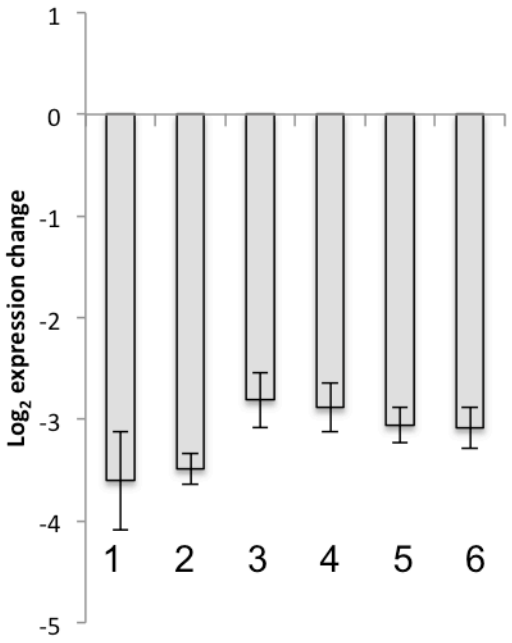


Figure 14: qRT-PCR expression analysis of NYR-v reporter expression. Detection of expression levels of YFP mRNA in six independent NYR-v (1-6) plants compared with a NYR-a control derived from leaf 4 RNA. Differential expression is shown as $\log_2$ fold change. Error bars display the standard deviation of results.

3.2.5 NYR-v undergoes gene silencing

To confirm a reduction in NYR expression, qRT-PCR was conducted to estimate the expression of the reporter (Fig.14). Six randomly selected NYR-v plants were tested showing an average -3.2 $\log_2$ change in expression compared with an NYR-a plant (Fig.14). qRT-PCR analysis utilised comparisonof expression with the constitutive expression of the housekeeping gene *gapdh*, selected over the widely used *actin* gene due to greater consistency of expression following in-house testing. These data show confirm a significant reduction in reporter expression in NYR-v plants.

3.2.6 Methylation analysis of NYR-v transgene

3.2.6.1 Methylation analysis by McrPCR DNA methylation has often been found to be responsible for silencing events in the functioning of gene and transgenes (McGinnis et al. 2006; Madzima et al. 2011; Matzke et al. 2009; He et al. 2011). Therefore, McrPCR was conducted in the UPN region, e35S promoter, YFP reporter of NYR-v and the unmethylated *Fertilisation-independent endosperm 2* (*Fie2*) control region (Fig.15). This allowed for analysis of the UPN region and the functional elements of the transgene, namely the e35S promoter and the YFP reporter. McrPCR analysis provides a basic view of the methylation state of a region, detecting

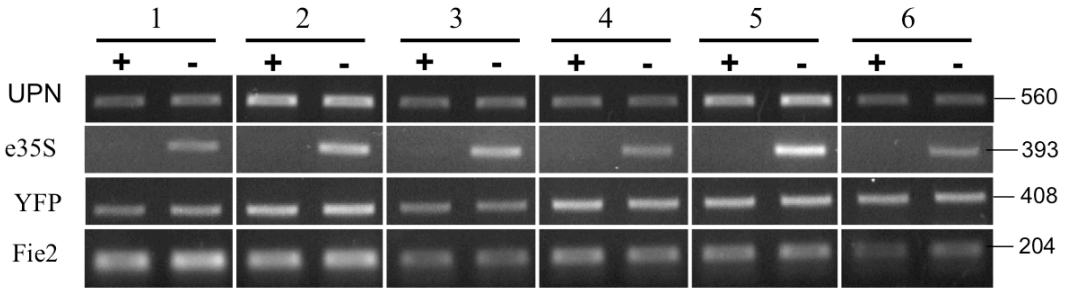


Figure 15: McrPCR methylation analysis of NYR-v. McrBC digestion and PCR amplification of the UPN region (position -1312 to -753) , e35S promoter (position -636 to -244), YFP (position +1589 to +1996) and *Fie2* (unmethylated control) of six NYR-v plants, gDNA derived from leaf 4 tissue. PCR product sizes (in bp) is indicated on right hand side of the panel.

the difference between heavy (hypermethylation) and light (hypomethylation) methylation only. An ideal comparison would have been NYR-v plants with NYR-a plants from separate transformation events, however the seed stocks were not avaliable. This analysis revealed that NYR-v is hypermethyated in the promoter and hypomethylated in the UPN region and YFP coding region.

3.2.6.2 Methylation analysis by bisulfite sequencing In order to get a detailed profile of the methylation status of the NYR-v promoter, bisulfite sequencing was utilised. Resulting data has been represented showing average cytosine methylation levels separated across different sequence contexts (Fig.16A) and methylation at individual cytosine positions (Fig.16B). Bisulfite treatment and sequencing was also conducted on the UPN region of NYR-v and the promoter of *Fie2*, both lacking methylation (Appendix Fig.68 and 69).

The promoter region of NYR-v has an average methylation percentage of 73.0%. This can be subdivided into 87.7%, 85.6% and 67.5% methylation of CG, CHG and CHH sequence contexts respectively. Observing the distribution of methylation across the promoter (Fig.16B) reveals a reduction in methylation at the 3' end. This region includes the *asf-1* (*activation sequence factor 1*), CAAT-like elements and TATA element, important in the binding of transcription factors, and also known sites for interference by DNA methylation (Fig.16C) (Kanazawa et al. 2007a). It is also a unique region in the promoter, not included within the tandem elements immediately upstream. The two CG positions within the *asf-1* element (positions -78 and -66) show 86.7% and 66.7% methylation and an average of 70.0% methylation including all cytosine positions. Additionally, the CAAT-like 1 element upstream of *asf-1* shows an average of 86.7% methylation, the later CAAT-like 2 shows 20.0% and CAAT-like 3 shows 26.7% methylation.

Furthermore, due to the variegated expression of the transgene the methylation status of individual clones were observed (15 clones in total) (Fig.17). Methylation of the promoter region ranged from

59.1% to 83.1%, 1 clone with 51-60%, 4 clones with 61-70%, 6 clones with 71-80% and 4 clones with 81-90% methylation. The *asf-1* element (only four cytosines) shows 1 clone with 0%, 1 clone with 25%, 2 clones with 50%, 7 clones with 75% and 4 clones with 100% methylation.

Collectively, these analyses show abundant DNA methylation at the NYR-v promoter only.

3.2.7 Treatment of NYR-v plants with inhibitors of DNA methylation

In order to determine if the presence of DNA methylation alone is responsible for NYR-v silencing, I conducted experiments using the DNA methylation inhibitors dihydroxypropyladenine (DHPA) (Koukalová et al. 2002) and Zebularine (Baubec et al. 2009). Both inhibitors have been shown to cause global reductions of DNA methylation across all sequence contexts. The roots of NYR-v seedlings were treated and subsequently screened for reactivation by confocal microscopy (Appendix Table.17). No reactivation was observed following the treatment. In addition, methylation analysis by McrPCR revealed that hypermethylation remained at the NYR-v promoter region and also at the Fie1 promoter, regardless of DHPA/Zebularine treatment (Fig.18).

Because the DHPA/Zebularine treatment of roots was unsuccessful in reactivating the NYR-v transgene, I imbibed seeds with both DNA methylation inhibitors prior to germination. This technique has recently been shown to be efficient in removing gene silencing in rice (Eun et al. 2012). In addition to DHPA and Zebularine, 5-azacytidine (5-Aza) (Eun et al. 2012) and Sulfamethazine (SMZ) (Zhang et al. 2012) were also utilised, all known to cause a global reduction in DNA methylation.

Seeds were imbibed for 16 hours at different concentrations of methylation inhibitors (Sulfamethazine, 50µM and 100µM; Zebularine, 40µM and 80µM; DHPA, 100µM and 200µM; 5-Aza, 5mM and 10mM). However, no reactivation was observed in roots following germination (Appendix Table.18).

McrPCR methylation analysis was conducted on the NYR-v promoter following these treatments (Fig.19). This analysis shows hypermethylation is still present in the promoter after treatment with Sulfamethazine, Zebularine and DHPA, however, a slight reduction in methylation is shown in some seeds treated with 5-Aza as shown by faint banding. In addition, the *Fertilisation-independent endosperm 1* (*Fie*1) promoter region tested also showed hypermethylation, despite all DNA methylation inhibition treatments.

Collectively, these experiments did not efficiently remove methylation from the NYR-v promoter as expected, and also did not result in NYR reactivation.

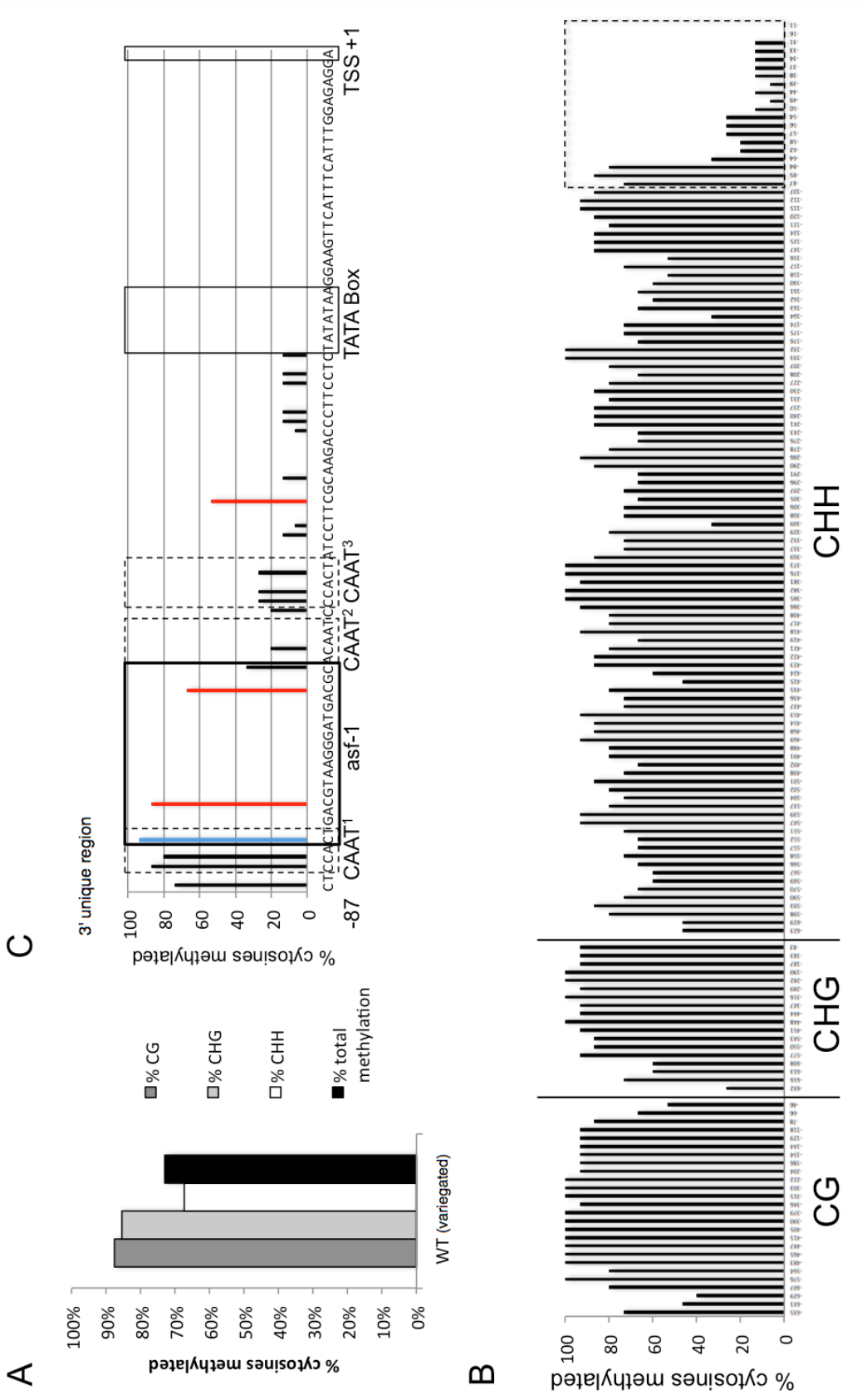


Figure 16: Bisulfite methylation analysis of NYR-v promoter region. PCR amplification and sequencing of bisulfite treated gDNA focusing on the NYR-v e35S promoter region (silenced) reports DNA methylation levels. The level of methylation is reported as the average percentage of total cytosines exhibiting methylation established from 15 independent clones for each data set from position -635 to -11 of NYR-v transgene, a total of 154 cytosines. Average methylation levels of (A) whole region, (B) individual cytosine positions, segregated into CG, CHG and CHH sequence contexts and (C) individual cytosine positions of the 3' unique region. CG positions (red), CHG (blue) and CHH (black).

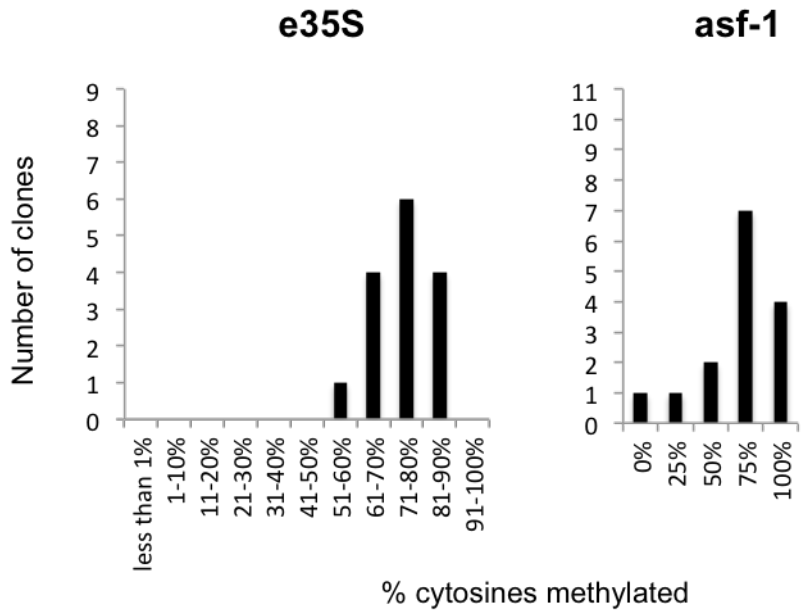


Figure 17: DNA methylation status of individual clones across the NYR-v e35S. The level of methylation for individual clones is reported as a percentage of cytosines exhibiting methylation in the NYR-v e35S promoter and the asf-1 element of the promoter (only four cytosines). 15 clones had been sequenced for each dataset.

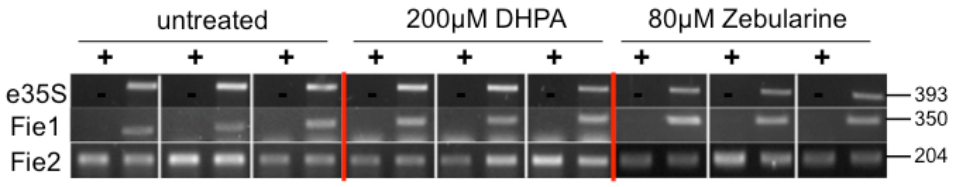


Figure 18: McrPCR methylation analysis of NYR-v roots treated with DNA methylation inhibitors. McrBC digestion and subsequent PCR amplification of the NYR-v e35S promoter, *Fie1* (methylated control) and *Fie2* (unmethylated control) following growth in water (untreated), DHPA or Zebularine (treated) of three individuals. PCR product sizes (in bp) is indicated on right hand side of the panel.

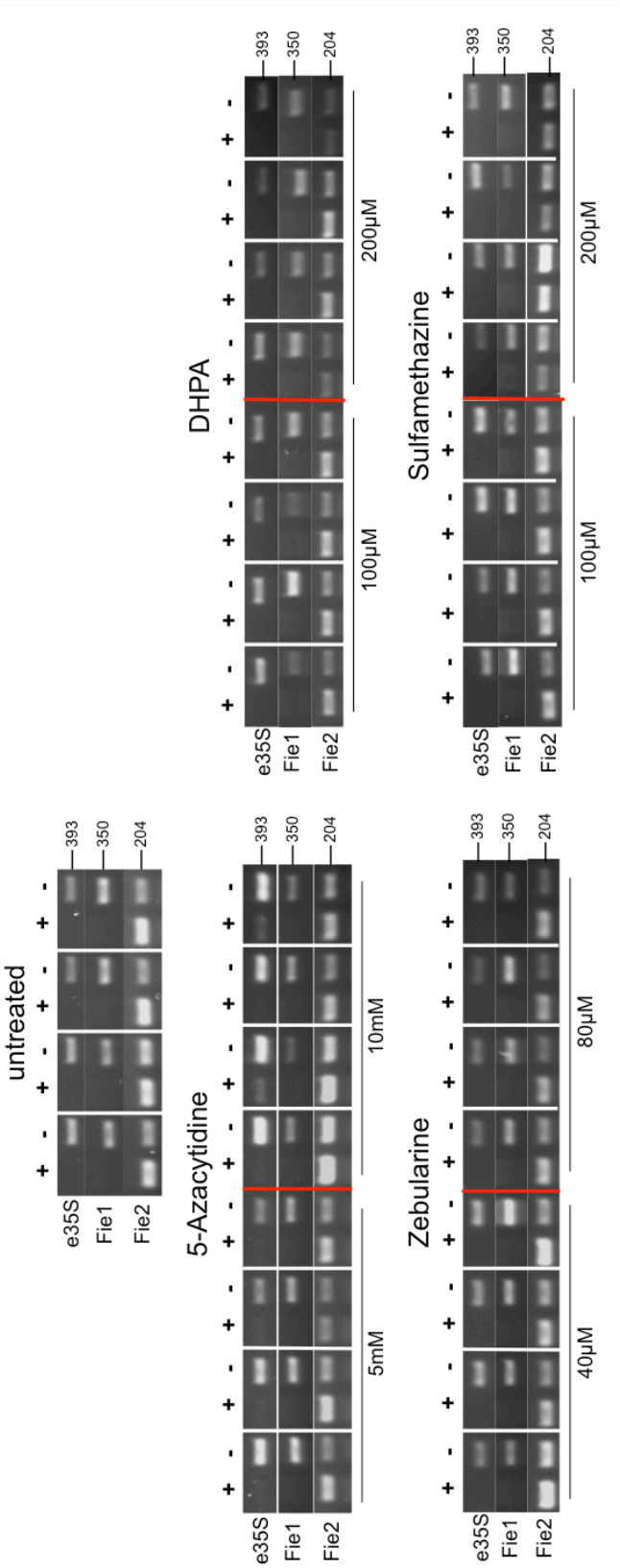


Figure 19: McrPCR methylation analysis of NYR-v roots treated with DNA methylation inhibitors during seed imbibing. McrBC digestion and PCR amplification of the NYR e35S promoter, *Fie1* (methylated control) and *Fie2* (unmethylated control) was conducted on four individuals each following imbibing seeds with DNA methylation inhibitors, 5-Azacytidine, DHPA, Zebularine and Sulfamethazine of varying concentration. PCR product sizes (in bp) is indicated on right hand side of panel.

3.2.8 Methylation analysis of NYR-v in vegetative and reproductive tissues

To determine if NYR-v DNA methylation patterns remained stable in differing tissues/organs, I conducted McrPCR methylation analysis. This analysis shows hypermethylation of the NYR-v promoter and hypomethylation of the UPN region and YFP in roots and leaves 3-12 tested (Fig.20). In addition, hypermethylation or the promoter and hypomethylation of the UPN region and YFP was also observed in ear and tassel tissue. Finally, analysis of pollen, embryos and endosperm revealed hypermethylation of the promoter in pollen and embryos, but hypomethylation in endosperm. The UPN and YFP regions for these reproductive tissues also shows hypomethylation.

Of specific relevance was endosperm tissue, which is known to have greatly reduced DNA methylation in *Arabidopsis* (Hsieh et al. 2009) and also similarly indicated in maize (Lauria et al. 2004). In maize this state is known to lead to the mobilisation/activation of TEs (Pan and Peterson 1988) due to relaxation of heterochromatin DNA. To test this hypothesis, I isolated seeds from NYR-v transgenic plants crossed with WT in addition to the other vegetative tissues tested. This analysis shows that hypomethylation is present in the UPN region and YFP of all tissues tested (as shown previously, see 3.2.6), and hypermethylation is present in the e35S promoter of all tissues except the endosperm, which shows hypomethylation (Fig.20).

These data show that the NYR-v promoter is differentially methylated in endosperm tissue.

3.2.9 Analysis of NYR-v expression in endosperm and embryo tissue

Hemizygous NYR-v transgenics were backcrossed with wild type plants in reciprocal orientations, and embryos and endosperm were harvested at 8 and 12 DAP. NYR-v reporter expression was analysed by epifluorescence microscopy of horizontally sectioned seeds, which exposed the endosperm and embryo tissues (Fig.21). At 8 DAP no expression of the reporter was observed in endosperm and embryo tissues (data not shown). At 12 DAP, of the 655 seeds observed from maternal transmission of NYR-v, 243 (37.1%) showed YFP expression in the endosperm tissue only (Fig.22). The paternal transmission of NYR-v showed only 5 (1.2%) seeds out of 403 observed showing expression in endosperm tissue (Appendix Tables.19 and 20 for a breakdown of individual crosses). Of each of the maternal and paternal crosses half of the seeds would be expected to not carry the transgene due to the genetics of the initial cross, these are recorded as without expression as the individual seeds were not genotyped.

These data suggest a parent-of-origin specific/imprinted expression of NYR-v in endosperm tissue.

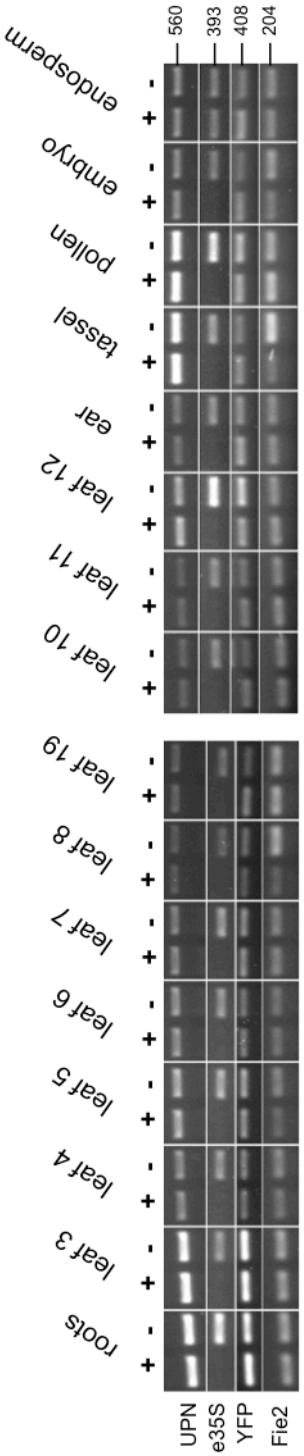


Figure 20: McrPCR methylation analysis of NYR-v in vegetative and reproductive tissues. McrBC digestion and PCR amplification of the NYR-v UPN region, e35S promoter, YFP and *Fie2* (unmethylated control) of vegetative and reproductive tissues. PCR product sizes (in bp) is indicated on right hand side of the panel.

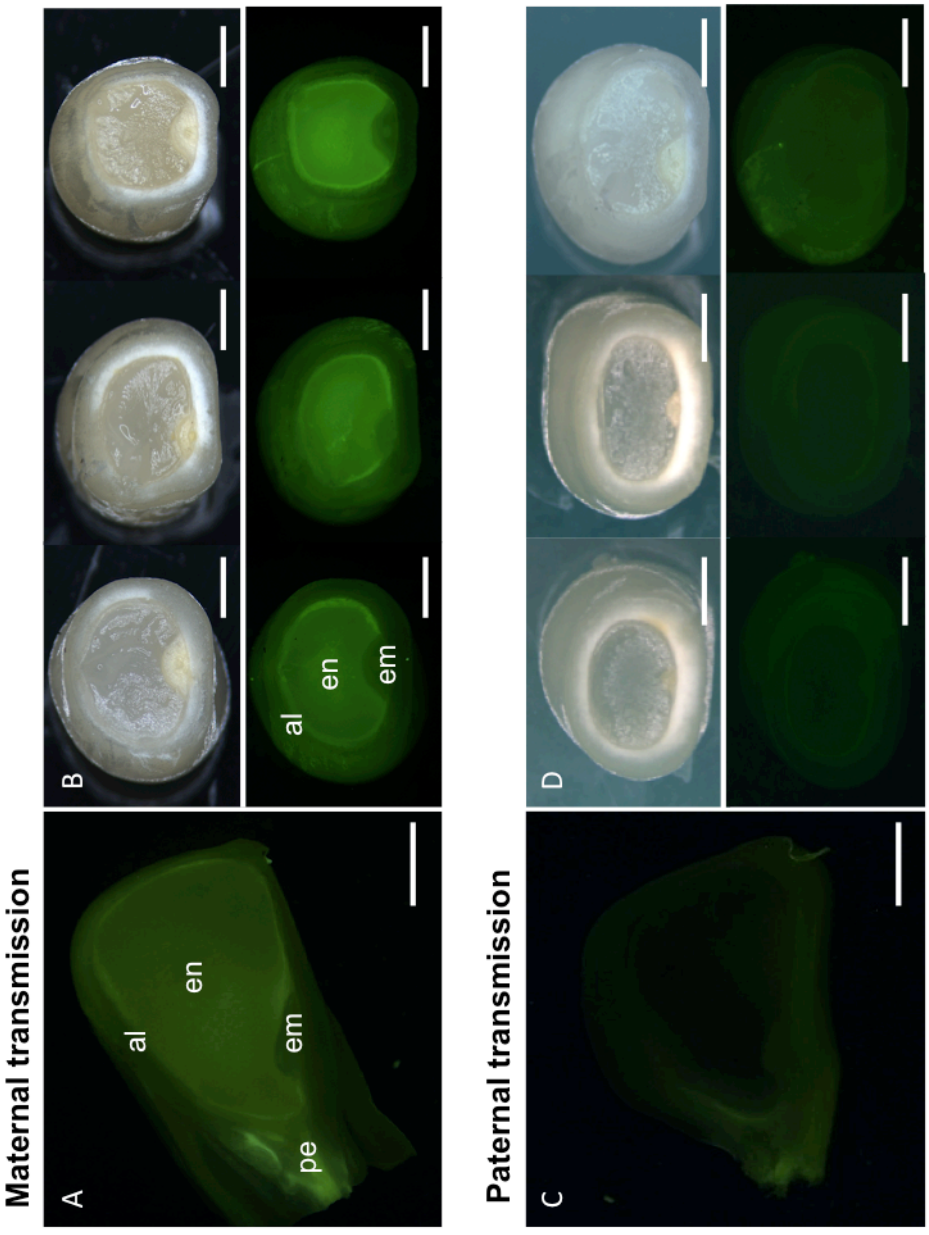


Figure 21: Epifluorescence microscopic images of YFP expression in maternally and paternally transmitted NYR-v. (A,B) Maternal transmission of NYR-v leads to active reporter expression in the endosperm tissue. (C,D) Paternal transmission fails to activate reporter expression. In both transmission orientations the embryo remains silent. Epifluorescence microscopic imaging, Horizontal cross section of 12 DAP seed (A,C) and transversal cross sections of 12 DAP seeds (B,D), top, light microscopy and bottom, epifluorescence. al, aleurone; en, endosperm; em, embryo; pe, pedicel. Scale bar = 2 mm.

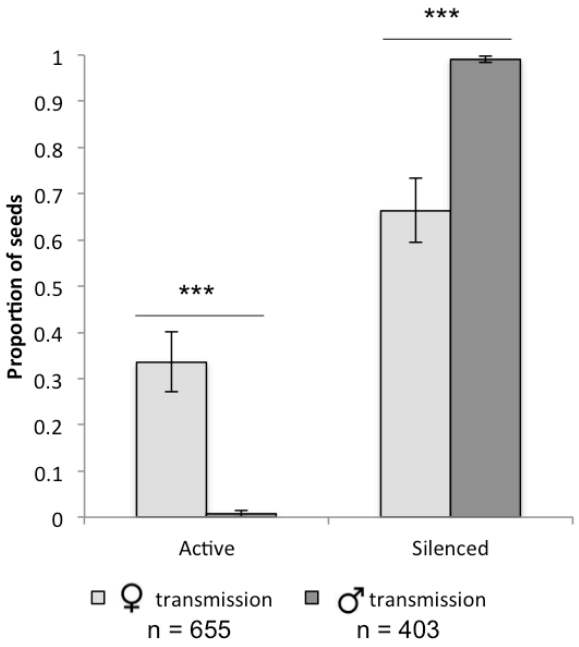


Figure 22: Proportion of seeds observed with active or silenced YFP expression in endosperm tissue from maternal and paternal transmission of NYR-v. Endosperm tissue (12 DAP) observed between maternal and paternal crossing of NYR-v transgene with WT plants. Analysis involved 655 (maternal transmission) and 403 (paternal transmission) seeds from eight independent crosses each. The asterisks marks the significant difference between active seeds produced from maternal and paternal cross ($p<0.001$), and the same for silent seeds as established by unpaired student's t-test.

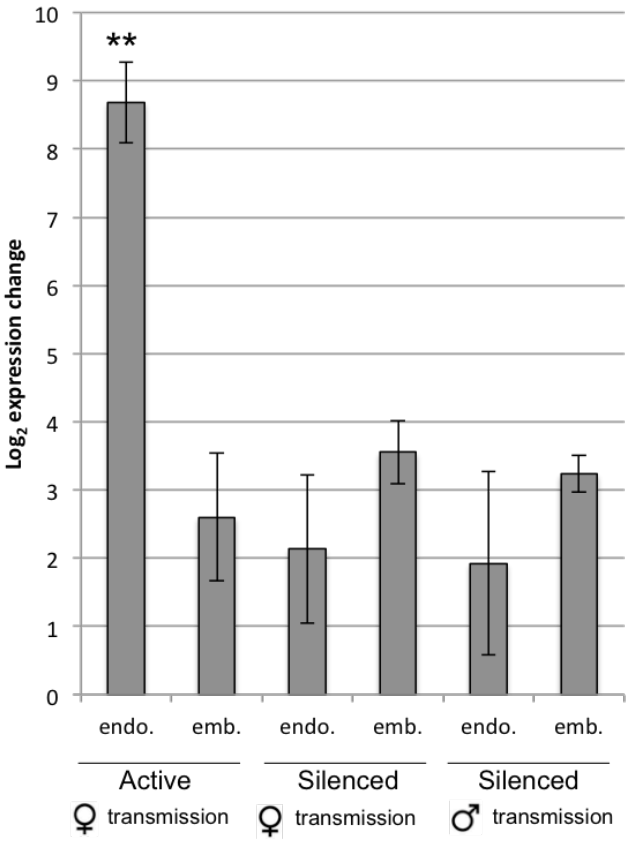


Figure 23: qRT-PCR expression analysis of YFP expression in maternally and paternally transmitted NYR-v. Detection of YFP expression in endosperm and embryo from maternally (active and silenced) and paternally (silenced) transmitted NYR-v transgene. Differential expression is shown as $\log_2$ fold change, compared to WT endosperm and embryo tissues. The two asterisks marks the significant difference between maternally transmitted active endosperm YFP expression and the other data points ($p<0.006$), as established by unpaired student's t-test.

3.2.10 Reporter expression of NYR-v in embryo and endosperm tissue

To assess the expression of the reporter in embryo and endosperm tissue qRT-PCR analysis was utilised (Fig.23). The expression of YFP in endosperms classified as active shows a 8.7 $\log_2$ change in expression compared with 2.1 and 1.9 from silenced maternal and paternally transmitted endosperm respectively. Embryos tested show a 2.6, 3.6 and 3.2 $\log_2$ change in expression for association with maternally transmitted active and silenced endosperm and paternally transmitted silenced endosperm respectively.

These data confirm the maternal specific expression of NYR-v in endosperm tissue.

3.2.11 Methylation analysis of NYR-v promoter in embryo and endosperm

3.2.11.1 Methylation analysis by McrPCR The methylation state of the NYR-v promoter was investigated by McrPCR methylation analysis (Fig.24). This analysis reveals that the promoter

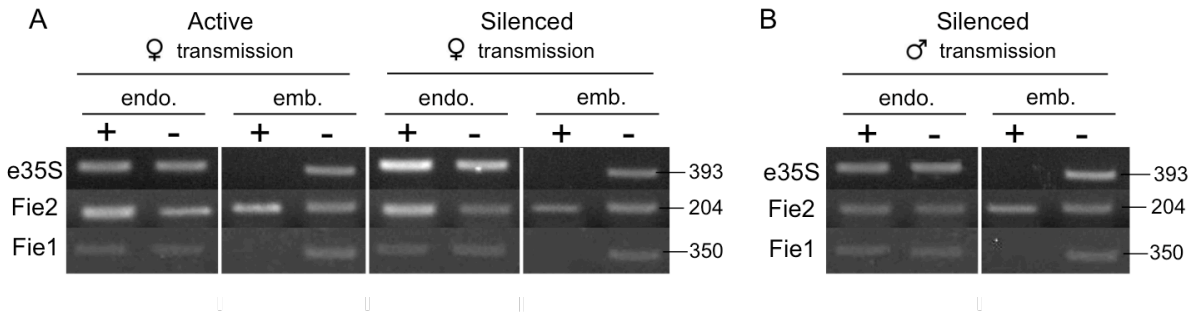


Figure 24: McrPCR methylation analysis of maternally and paternally transmitted NYR-v. McrBC digestion and PCR amplification of NYR-v e35S promoter, *Fie2* (unmethylated control) and *Fie1* (methylated endogenous control) of 25 pooled endosperm and embryo tissue in (A) maternally transmitted NYR-v, active and silenced states and (B) paternally transmitted NYR-v silenced state. endo., endosperm; emb., embryo. PCR product sizes (in bp) are marked on the right of the panel.

is hypomethylated in endosperm and hypermethylated in embryos, regardless of parental transmission. Also, analysis of the imprinted demethylated region (DMR) of *Fie1* showed hypomethylation in endosperm and hypermethylation in embryo tissue (Gutiérrez-Marcos et al. 2006).

3.2.11.2 Methylation analysis by bisulfite sequencing Bisulfite sequencing of the NYR-v promoter region was conducted to get a detailed DNA methylation profile at the single nucleotide level and averaged across the promoter region (Fig.25).

The average levels of methylation in the promoter were 47.5% (CG, 69.2%; CHG, 54.6%; CHH, 41.2%), 38.7% (CG, 50.6%; CHG, 41.7%; CHH, 35.5%) and 43.9% (CG, 53.2%; CHG, 56.9%; CHH, 40.6%) in maternally transmitted active and silenced endosperm and paternally transmitted silenced endosperm respectively. Silenced maternally transmitted endosperm tissue was significantly reduced in methylation compared with the other two endosperm samples, however this was a very small difference. Associated embryos showed 61.0% (CG, 89.1%; CHG, 75.9%; CHH, 52.0%), 63.3% (CG, 96.8%; CHG, 79.6%; CHH, 52.7%) and 58.8% (CG, 90.4%; CHG, 76.7%; CHH, 48.4%) average methylation of maternally transmitted active and silenced endosperm and paternally transmitted silenced endosperm respectively. Comparatively, the average methylation of the NYR-v in leaf material was shown to be 73.0%.

Segregating the different cytosine methylation contexts shows a reduction in CHH methylation at the 3' unique region of the promoter (Fig.26). In addition, the methylation profile of the 3' unique region shows the *asf-1* region with 37.5% (CG positions -78 and -66, 50.0% and 33.3% methylated), 58.3% (50.0% and 66.7% methylated) and 29.2% (16.7% and 33.3% methylated) average methylation in maternally transmitted active and silenced endosperm and paternally transmitted

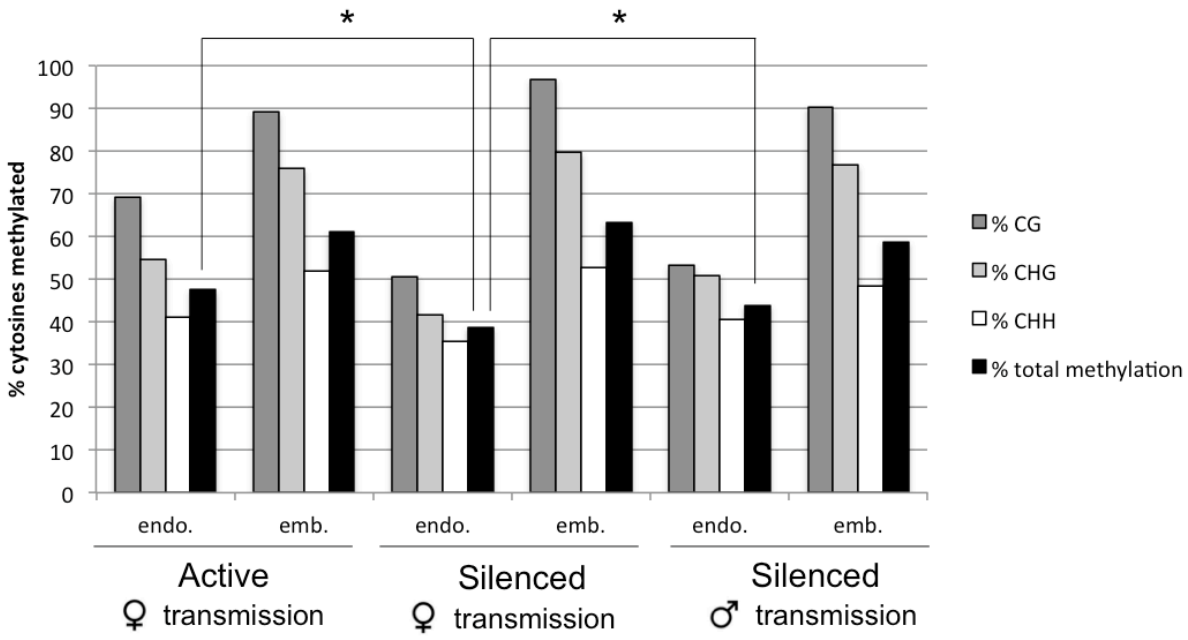


Figure 25: Bisulfite methylation analysis of the promoter region in maternally and paternally transmitted NYR-v. PCR amplification and sequencing of e35S promoter region from bisulfite converted gDNA of maternally and paternally transmitted NYR-v extracted from pooled endosperm and embryo tissues (25 each). Active YFP expressing and silenced tissues were analysed from maternal transmission and only silenced tissue from paternal transmission. The level of methylation is reported as the average percentage of total cytosines exhibiting methylation established from six independent clones for each data set, from position -635 to -11 of NYR-v transgene, a total of 154 cytosines. The asterisk marks the significant difference between endosperm of active maternally/silenced paternally transmitted NYR-v and silenced maternally transmitted NYR-v ($p>0.05$).

silenced endosperm respectively (Fig.27). The embryos show the *asf-1* region as 70.8% (CG positions -78 and -66, 100.0% and 83.3% methylated), 100.0% (100.0% and 100.0% methylated) and 87.5% (83.3% and 100.0% methylated) average methylation of maternally transmitted active and silenced endosperm and paternally transmitted silenced endosperm respectively.

Average methylation of the unique 3' promoter region as a whole in (position -87 to -11) shows 27.8%, 29.2% and 17.4% (significantly different between silenced samples) in maternally transmitted active and silenced endosperm and paternally transmitted silenced endosperm respectively. The average methylation was shown as 45.8%, 53.5% and 49.3% in embryo of maternally transmitted active and silenced endosperm and paternally transmitted silenced endosperm respectively (full methylation values can be found in Table.21 in the Appendix).

Despite differing reporter expression levels, no significant differences are detected in the total DNA methylation of the promoter region of active or silenced NYR-v endosperm.

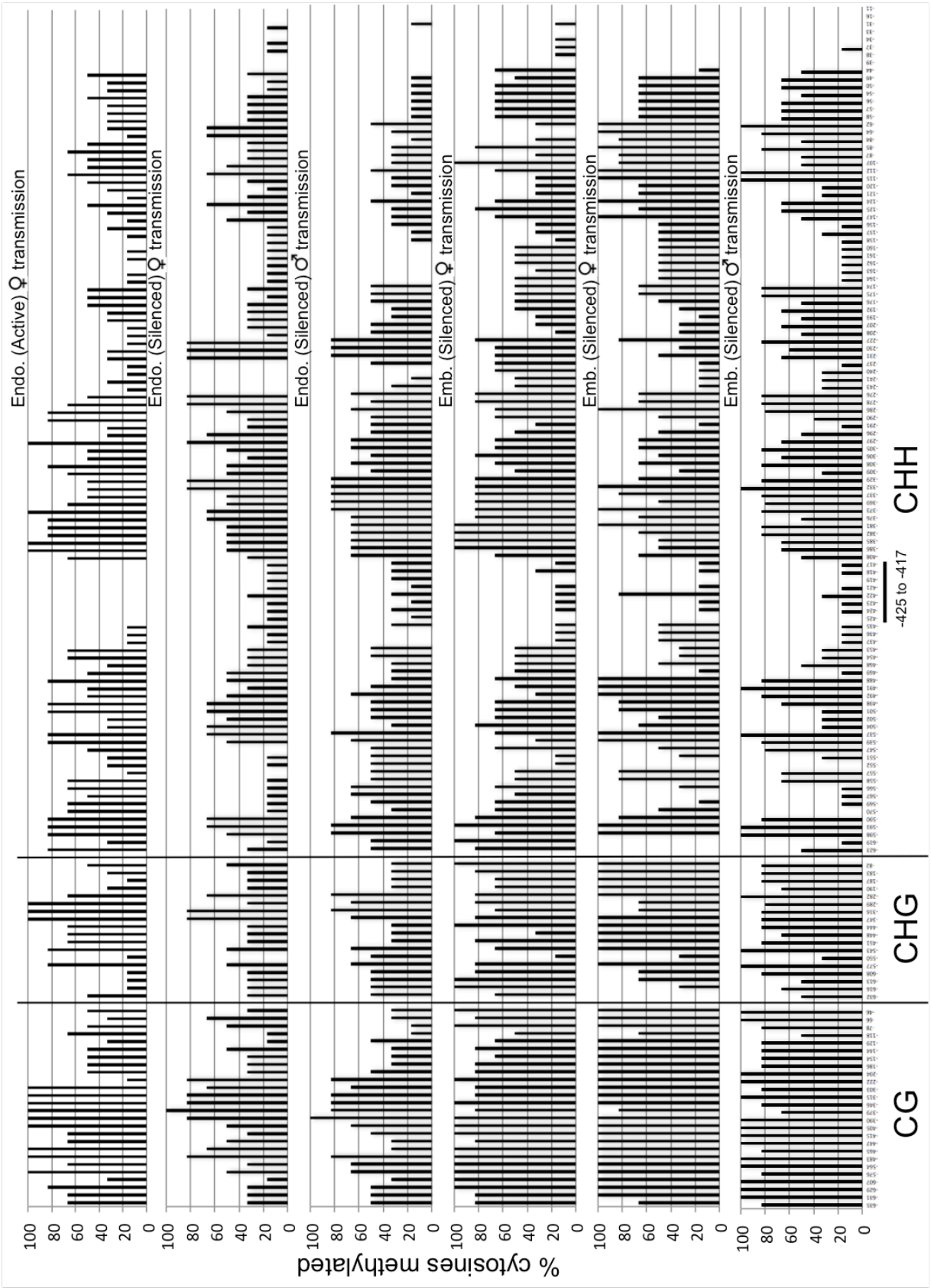


Figure 26: Methylation analysis of e35S promoter showing CG, CHG and CHH methylation in maternally and paternally transmitted NYR-v endosperm and embryo tissue. Bisulfite analysis of DNA methylation in the NYR-v e35S promoter of maternally transmitted NYR-v, 'active' and 'silenced' states and paternally transmitted NYR-v 'silenced' state. The level of methylation is reported as the percentage of individual cytosine locations exhibiting methylation, established from 6 independent clones, spanning -635 to -11 of NYR transgene (cytosine positions marked on the x-axis), a total of 154 cytosines.

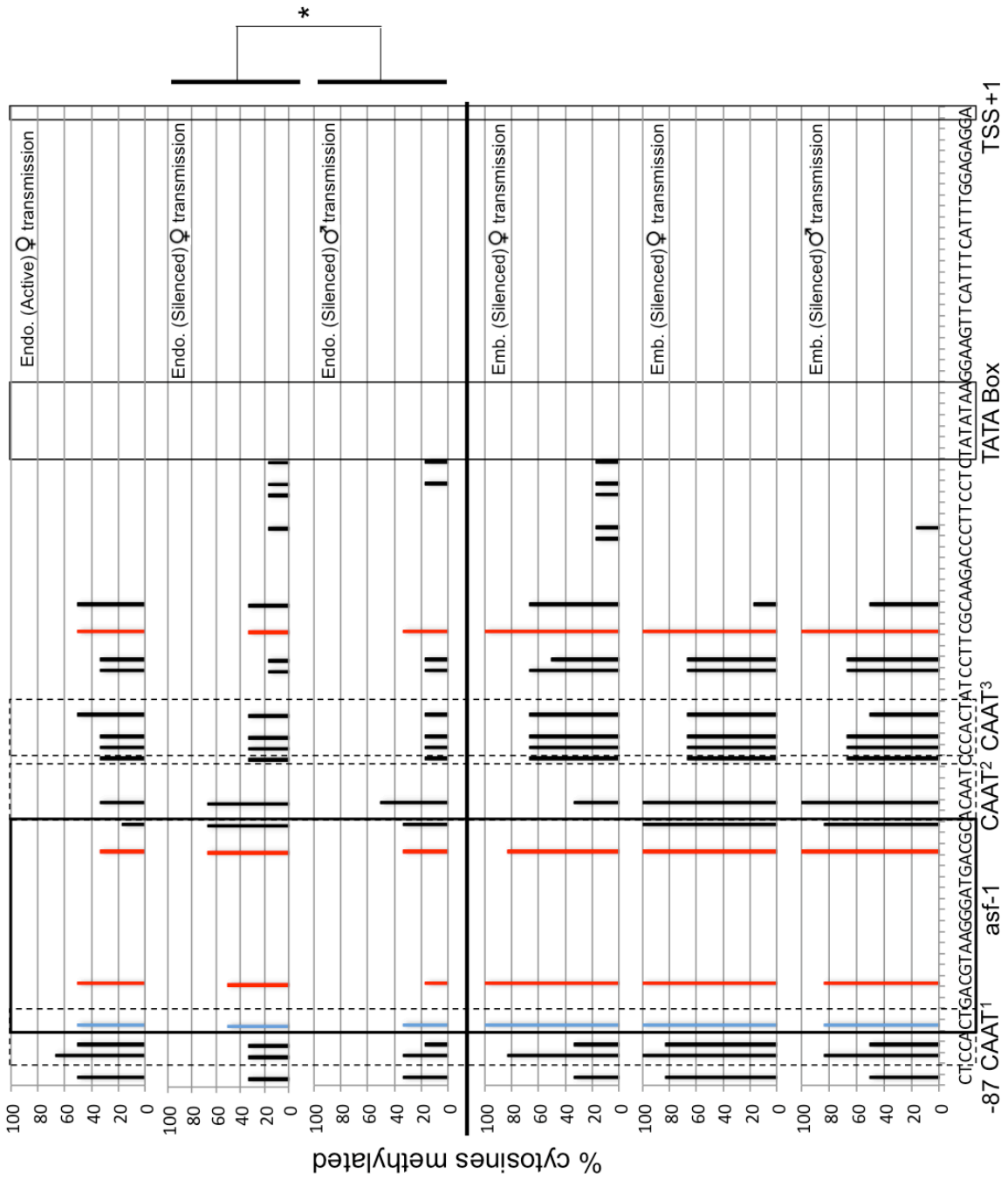


Figure 27: DNA Methylation profile of NYR-v 3' promoter region in maternally and paternally transmitted endosperm and embryos. Methylation levels between position -87 and +1 of the NYR e35S promoter displaying CAAT-like elements, asf-1 element, TATA box and TSS (+1) of maternally transmitted NYR-v, 'active' and 'silenced' states and paternally transmitted NYR-v, 'silenced' state, with red representing CG, blue as CHG and black as CHH methylation. The level of methylation is reported as the percentage of individual cytosine locations exhibiting methylation, established from 6 independent clones (position number marked on the x-axis). The asterisks denote the significant difference between average levels of methylation ($p<0.05$) of the data sets as established by Wilcoxon two sample rank test.

3.3 Discussion

3.3.1 Regulation of NYR-v expression

Variegated expression or silencing of transgenes in plants has long been attributed to epigenetic regulation (TGS) (Park et al. 1996; Meyer and Saedler 1996; Madzima et al. 2011) and PTGS (Ingelbrecht et al. 1994; Elmayan and Vaucheret 1996). qRT-PCR expression analysis of NYR-v individuals confirms a reduction of YFP mRNA as observed in roots and leaves when compared with NYR-a. These observations show the NYR-v transformation event presents a non-tissue specific, stochastic/dynamic process of silencing. Root epidermal cells (also evident in leaves) reveal that once an active or silent state is achieved it often persists following mitotic divisions, as shown by lines or sectors of the same expression state (longitudinal) originating from the proximal root meristem. This suggests that an epigenetic state is somatically stable but can also be reversible, and has been observed in *FLC:GU*S fusions in *Arabidopsis* (Angel et al. 2011). Previous studies propose that variegated expression of transgenes is linked to inefficient silencing of their targets (Madzima et al. 2011).

Silencing mechanisms are thought to serve as defence against the introduction of foreign DNA such as transgenes, TEs and other invasive (viral) elements (Matzke and Matzke 1998). However, the modes of action against these elements are not well understood. A variety of situations can result in transgene silencing; multiple transgene insertions, transgene homozygosity, repeat sequences, insertion location, transcriptional intensity and insertion location, of which both PTGS and TGS mechanisms are implicated (Fagard and Vaucheret 2000).

3.3.2 NYR-v structure and position

Sequencing of NYR-v showed that the transgene has not undergone visible rearrangements. Therefore, any changes in expression could be attributed to epigenetic differences or silencing actions. The transgene copy number was also determined by restriction enzyme digestion and Southern Hybridisation analysis, indicating a single insertion in the maize genome. Although, single transgene copies have also been shown to be vulnerable to silencing, it is less likely to occur (Pröls and Meyer 1992). Furthermore, the promoter utilised is known to attract TGS (Meyer et al. 1994; Assaad et al. 1993; Khaitová et al. 2011) and also PTGS (Velten et al. 2012). The high amount of mRNA transcribed by this promoter is thought to trip a threshold, identifying the region to the genome as a potentially harmful, therefore requiring silencing (Fagard and Vaucheret 2000). It has been previously shown that aberrant RNAs generated from increased amounts of mRNA in

the cytoplasms can induce PTGS of both transgenes and endogenous genes (Linder and Owttrim 2009). This is used by plants to combat virual infections, repressing transcription, however, it also acts on non-viral elements. It is thought that high levels of expression can induce PTGS, known as the 'RNA threshold model' (Lindbo et al. 1993). This proposes that once transcription exceeds a threshold level, the PTGS pathway is initiated, degrading the targeted mRNA. The level of transcription required to initiate this response is unclear, which is complicated by PTGS of (trans)genes without high expression (Stam et al. 1997; Schubert et al. 2004). Furthermore, a more recent study has shown that silencing can be triggered if transcription surpasses a gene-specific threshold (Schubert et al. 2004). This investigation observed the expression levels of different transgenes, varying the copy number. The expression of a GFP transgene driven by a 35S promoter showed high expression in plants with four copies, but reduced expression in plants possessing more than five copies (Schubert et al. 2004). This threshold for silencing was found at different copy numbers for different transgenes tested, with a GUS reporter transgene requiring only three copies to show reduced expression. A variegated expression may therefore be attributed to levels of expression straddling this threshold, and may be the cause of NYR-v expression. This change in expression could be facilitated by natural variation in expression or perhaps dynamic chromatin processes inhibiting access to the transgene, dependent on position.

On the other hand, sequence homology is also known to attract silencing, recruiting epigenetic machinery to sites of repetitive DNA, at the same location or distant regions of homology (Assaad et al. 1993; Park et al. 1996; Meyer and Saedler 1996). The NYR-v promoter possess a tandem element of the first ~300 bp, which may therefore attract epigenetic silencing. The presence of DNA methylation mediated by siRNAs directed to the promoter region would result in epigenetic TGS. This mechanism acting on NYR-v is discussed more in 3.3.3.

The insertion location of NYR is also of particular interest (as previously stated), as chromosomal position has often been thought to be involved in expression efficiency in plants (Singh et al. 2008). Insertion into a region of dense heterochromatin/repetitive sequence would suggest that there is a probability of physical interference of transcription factor interactions, potentially silencing (Pröls and Meyer 1992; Tartof et al. 1984). The most well known study of which shows PEV in *Drosophila* (Tartof et al. 1984), but studies have also been conducted in plants claiming the same effects (Matzke and Matzke 1998; Pröls and Meyer 1992). However, other studies argue the opposite, (Nagaya et al. 2005; De Buck et al. 2004) stating that these events are most likely RdDM acting on repetitive elements and that position has little or no effect on expression levels. Intriguingly, the NYR-v transgene is flanked by a *gypsy*-like LTR TE. This type of TE is known

to comprise 0.62% of the maize genome (Vicient 2010) and comes as no surprise, as 85% of the maize genome is thought to be comprised of TEs (Schnable et al. 2009), with 95% of those being comprised of LTRs (Haberer et al. 2005). In plants TEs are largely methylated correlating with transcriptional silencing, however, a study in maize has found that *gypsy*-like TEs show high levels of transcriptional activity (Vicient 2010). This suggests that it is more likely that the flanking region possess a relaxed chromatin state, however, in the event of a repressive heterochromatin environment studies have shown epigenetic marks spreading into neighbouring genic sequences (Eichten et al. 2012). The transcriptional and methylation profiles of the upstream flanking region may provide greater information but were not explored.

Only a single transformation event produced NYR with a variegated phenotype, and currently the only known difference is the insertion location of the transgene. Therefore, it is possible that the upstream endogenous DNA may influence NYR-v expression by chromatin spreading, causing a positional effect.

3.3.3 The NYR-v promoter is hypermethylated

Transgene silencing has often been associated with the presence of DNA methylation in both TGS and PTGS events (Fagard and Vaucheret 2000). Generally, TGS has been identified with methylation targeted to the promoter region (Meyer et al. 1993; Park et al. 1996; Madzima et al. 2011) and PTGS derived methylation within coding regions (Fagard and Vaucheret 2000; Paszkowski and Whitham 2001). Methylation analysis of NYR-v revealed no methylation in the YFP reporter or H2B sequence (data not shown), however, high levels of methylation were found at the promoter. Further analysis concluded that all sequence contexts of methylation were represented, including *de novo* CHH, which is associated with the action of RdDM pathway (He et al. 2011). The highest levels of methylation observed was found at the tandem repeat elements of the promoter, suggesting that the repetitive sequences has attracted *de novo* DNA methylation. The 3' unique region of the NYR-v promoter, which possesses the transcription factor binding site *asf-1* showed lower levels of DNA methylation. The *asf-1* sequence is an essential factor in transcriptional efficiency of the 35S promoter and a mutation within its 21 bp sequence has been found to significantly down regulate promoter function (~50%) (Lam et al. 1989). In addition, the presence of DNA methylation at CG positions of *asf-1* is also associated with transcriptional silencing of the promoter (Meyer et al. 1994; Kanazawa et al. 2007a). NYR-v presents high levels of methylation at both CG positions, in addition to other asymmetric sites within *asf-1*, which may be in part responsible for silencing. However, other elements important for transcription factor binding, CAAT-like elements, TATA

element and the Transcriptional Start Site (TSS) are present in the promoter (Ow et al. 1987) have also been found to be sensitive to the presence of DNA methylation (Kanazawa et al. 2007a). The NYR-v promoter also shows methylation of all sequence contexts across these binding elements, which may also contribute to silencing.

Furthermore, a mix of methylation profiles was observed in clones analysed for bisulfite methylation analysis, ranging between 59.1% to 83.1% total methylation across the promoter region. 11 out of 15 clones also showed 75% or more methylation at the *asf-1* region. Therefore, variegation of the transgene may be as a result of differential methylation of individual cells, shown by the mixed profiles. Methylation studies of the 35S promoter showing variable/partial expression of total cells have also demonstrated such a mix (Brabbs et al. 2013). This study shows that mutation of MORC6 (microrchidia family ATPase) results in the variable expression of the transgene, leading to obvious DNA methylation increases within a proportion of the clones used for bisulfite sequencing analysis. Furthermore, separation of active and silenced cells by Fluorescence-Activated Cell Sorting (FACS) and methylation analysis revealed uniform profiles between cells of the same expression levels (Brabbs et al. 2013). The majority of NYR-v clones showing methylation within the *asf-1* region and promoter as a whole may therefore be responsible for repression of promoter function in the majority of cells. Without methylation analysis of other transformation events generating NYR-a it remains unknown whether methylation is exclusive to NYR-v.

On these grounds the removal of methylation was expected to reactivate NYR-v, implicating DNA methylation as the direct cause of silencing. However, a reduction in NYR-v promoter methylation was only observed with 5-Aza treatment of a proportion of seeds and did not restore reporter expression. Bisulfite sequencing would be best employed to determine the methylation profile of seeds that have undergone 5-Aza treatment. Currently a slight loss of methylation is interpreted from the McrPCR analysis due to the presence of faint banding is not a strong indication of the removal of methylation. *Fie1* controls also failed to reciprocate this result, remaining methylated.

3.3.4 Imprinted expression of NYR-v

Reciprocal crosses carried out with NYR-v showed that maternal transmission was associated with active endosperm expression (37.1%) compared to paternal transmission (1.2%). Epifluorescent microscopy was backed up by qRT-PCR data showing the greater amount of YFP expression.

To date, no entirely synthetic transgene has been shown to be imprinted in plants, although, it has been observed in mammals, and is associated with differential methylation patterns dependent on parent-of-origin inheritance (Chaillet et al. 1991; Reik et al. 1987). Imprinted expression has,

however, been observed in transgenes incorporating endogenous imprinted sequences in plants (Kinoshita et al. 2004). Currently only 11 genes have been identified showing imprinted expression in maize (Raissig et al. 2011). Parental genes are epigenetically coded prior to fertilisation in the gametes or in the seed tissues following fertilisation, usually observed as differing methylation profiles. The parental orientation of transmission of coded epialleles can then effect their expression, most often in plants observed in the endosperm. A large number of candidate imprinted genes have been identified in maize endosperm tissue by analysis of transcriptional differences and the identification of differential transmission of DNA methylation patterns contributed by parental alleles (Lu et al. 2013; Waters et al. 2011; Gehring et al. 2009; Haun and Springer 2008; Kinoshita et al. 2004).

In maize and other flowering plants genome-wide demethylation of repetitive DNA occurs in the endosperm. This decrease in methylation has been utilised to aid in the identification of imprinted genes, which can show lower levels of DNA methylation compared to embryo tissue (Gutiérrez-Marcos et al. 2006; Lauria et al. 2004; Gehring et al. 2009). This difference in methylation can be seen in NYR-v, comparing endosperm and embryo methylation profiles of all cross orientations tested. In *Arabidopsis* genome-wide demethylation of the central cell occurs prior to fertilisation through the action of DME (DEMETER) (Hsieh et al. 2009; Gehring et al. 2006). However, DME function is not conserved in maize, therefore, the current theory is that demethylation of maternal alleles occurs in endosperm when the paternal alleles remain methylated.

My results show that methylation patterns of maternally and paternally transmitted NYR-v do not differ significantly in the endosperm. Furthermore, there is no significant difference between NYR-v promoter methylation between endosperm classified as active and silent in the maternal transmission. Therefore, there are no differing epialleles of the NYR-v promoter region to explain the difference in expression. However, methylation analysis of *Mee1*, the only gene imprinted in embryo tissue, shows a changing methylation profile dependent on the stage of seed development (Jahnke and Scholten 2009). The results state that DNA methylation does not always correlate with transcription, but instead is replaced by dynamic methylation profiles. But, despite dynamic methylation changes in the embryo, differing endosperm methylation profiles were still observed (Jahnke and Scholten 2009). The uniform reduced methylation state shown by the NYR-v promoter in endosperm tissue of both active and silence expression suggests that methylation does not correlate with expression in endosperm tissue. This could be the result of (i) histone modifications or (ii) undetected patterns of DNA methylation at other locations.

(i) Almost all current maize imprinted genes have shown differing epialleles in the endosperm

and a number have also shown association of specific histone modifications. A study conducted in *Arabidopsis* investigates the histone modifications present at *Mez1*, *ZmFie1* and *Nrp1*, all of which are imprinted in the maize endosperm and have differing methylation patterns (Haun and Springer 2008). Repressive histone modifications, such as H3K27me2 and me3 can be found in alleles silenced by paternal transmission and histone modifications promoting active transcription, such as, acetylation of H3 and H4 or H3K4me2 were observed in active maternal transmission of alleles (Haun and Springer 2008). Differing histone modifications may therefore provide the epigenetic variation alone for imprinting, although it has not been observed without different DNA methylation profiles.

(ii) Imprinting has also been found reliant on DNA methylation at unexpected sites. Work conducted in *Arabidopsis* found that the imprinted expression of PHERES1 was dependent on methylation at a downstream location (Makarevich et al. 2008). Methylation analysis of NYR-v was only conducted at the e35S promoter, therefore, the methylation profile of adjacent regions may play a role in controlling expression. The identification of new imprinted genes has also led to the hypothesis of an effect of proximity of TEs to imprinted genes (Mosher and Melnyk 2010; Gehring et al. 2009). The genome-wide reduction in DNA methylation at TEs in the endosperm suggests that there may be involvement with imprinted loci. As previously stated, flanking NYR-v is a *gypsy*-like LTR TE which may have a role in silencing, however, the role of TEs and repetitive elements in imprinting is still unclear.

The proportion of seeds showing active expression in the maternal cross orientation was also only 37%, with 50% of seeds expected to carry the NYR-v transgene. The same stochastic process that effects expression in vegetative tissue may also effect endosperm expression. On the other hand, this result may be the result of the development stage of the seeds observed. At 8 DAP, no endosperm displayed active levels of reporter expression from either cross orientation. Seeds derived from paternal transmission of NYR-v were also left for up to 16 DAP to ensure that active endosperms were not detected to the same levels at 12 DAP (data not shown). A later time point for maternally transmitted NYR-v may also have been beneficial to observe the proportion of active expressing seeds.

Alternatively, it could be argued that high expression of NYR-v from the maternal transmission is a result of a dosage dependent effect, a classic argument when applied to gene imprinting in the triploid endosperm of plants. However, this result would suggest that all seeds present uniform expression within each cross orientation, but as previously discussed, a proportion of maternally transmitted NYR-v seeds showed no expression.

Endosperm expression of NYR-v may be the product of imprinting, however, the controlling factor does not appear to rely on DNA methylation at the promoter. Therefore, the expression profile observed may be the result of unknown histone modifications or DNA methylation at different locations/adjacent endogenous regions.

3.3.5 Summary

In summary, the initiation of NYR-v silencing appears to involve DNA methylation mainly located at the promoter tandem repeat elements, possibly as part of TGS or homology-based silencing. Methylation spreading to the *asf-1* element and other functional binding elements of the promoter (differing between clones observed) may be implicated in the dynamic silencing, thus resulting in a variegated phenotype. Alternatively, PTGS could be the driving force behind silencing, however, the detection of a single copy number insertion of NYR-v would indicate that high expression levels required to exceed the endogenous gene threshold would not be achieved. Finally, insertion location may also have a role in silencing due to only a single transformation event inducing this expression profile.

My data has also revealed the imprinted expression of NYR-v in the endosperm. However, the DNA methylation profiles observed do not reflect traditional imprinted gene, leading the cause to be disputed.

4 The Role of RdDM in NYR-v Silencing

4.1 Introduction

4.1.1 Chapter aims

The previous chapter has identified the presence of DNA methylation at the NYR-v promoter, suggesting that the RdDM pathway is responsible for silencing. Previous experiments conducted in an effort to disrupt methylation through treatments with DNA methylation inhibitors failed to reactivate the transgene or reduce methylation at the control region. Therefore in order to determine if the RdDM pathway is involved in establishing methylation, small RNA sequencing was conducted. This analysis can conclude if there are small RNAs present within the NYR-v transgene responsible for the direction of methylation. Following this, mutants of the RdDM pathway were introgressed with NYR-v to remove methylation in an attempt to release silencing, determining the mechanisms utilised by NYR-v to establish and maintain silencing. Observation of reporter expression in the progeny can then determine if variegation has been replaced by unimpeded expression.

- Identify if RdDM mediated small RNAs are involved in silencing of NYR-v
- Determine if NYR-v silencing can be released through introgression with RdDM mutants.
- Analyse the effect of RdDM mutations on NYR-v methylation
- Lastly, following from the potential reactivation of NYR-v, if silencing can be efficiently established through restoration of the RdDM pathway.

The next section describes how small RNA analysis can be conducted with the use of next generation sequencing technologies followed by the mutants which can be utilised in maize for the study of RdDM facilitated silencing.

4.1.2 Small RNA analysis by next generation sequencing

Small RNAs are responsible for directing methylation, most notably the action of 21-24 nt siRNA size class (Mette et al. 2000). Initial studies of small RNAs used northern analysis in the detection of specific sequences (Hamilton and Baulcombe 1999), which has become a common-place technique. However, detection of low expressing small RNAs requires a high abundance of starting material, the protocol can be labour intensive and data is only generated on a single transcript. In order to produce small RNA transcriptome profiles, next generation sequencing technologies were developed. The first whole genome study using this technology was carried out in *Arabidopsis*,

investigating the transcriptome with Massively Parallel Signature Sequencing (MPSS) (Meyers et al. 2004) and shortly after, specifically small RNAs (Lu et al. 2005). DNA is bound to beads and transferred to a flow cell for a fluorescence-based detection of newly generated DNA strands (Brenner et al. 2000). Sequencing data generated can then be annotated back to homologous regions of a reference genome, showing the location and abundance of small RNAs. However, MPSS technologies were limited to a sequence length of 17-20 nt, short of the now known 21-24 nt siRNA size class. Subsequent technologies such as pyrosequencing (454 Life Sciences) addressed this, allowing for over 100 nt, but sacrificing the depth of sequencing (Nobuta et al. 2010). Current technologies combine very good sequence depth and increased sequence length detection, such as the sequencing-by-synthesis (SBS) system offered by Illumina. This technology has allowed for small RNA profiles to be generated with immense detail, detecting even minute expression (Nobuta et al. 2008). These studies, among numerous others have solidified the advantages of next generation sequencing of small RNAs over other techniques such as microarray and PCR-based analysis (Schmittgen et al. 2004-Humans, Krichevsky et al. 2003-Rats).

4.1.3 RdDM mutants used in the study of TGS

The study of the functional components of the RdDM pathway and its operations has been facilitated by mutations perturbing the pathway. In both *Arabidopsis* and maize they have proved invaluable in demonstrating direction of DNA methylation and through this the dissection of paramutation (McGinnis et al. 2006; Sidorenko et al. 2009; Hollick and Chandler 2001), small RNA movement (Molnar et al. 2010) and TGS (Madzima et al. 2011; McGinnis et al. 2006) to name a few. Table.11 provides a summary of maize RdDM mutants (all used in analysis later in the chapter), those that affect the accumulation of siRNA and those that have shown to interfere with TGS and hypothesisied to be involved in mediating silencing.

Genetic screening by the *b1* and *pl1* systems have allowed for the identification of mutants involved in paramutation, these include the characterisation of *mop1* and *mop2* mutants (using the *b1* allele) (Dorweiler et al. 2000; Sidorenko et al. 2009) and *rmr1*, *rmr2* and *rmr6* mutants (using the *pl1* allele) (Hale et al. 2007; Erhard et al. 2009; Hollick and Chandler 2001). Furthermore, *rmr1*, *mop1* and *rmr2* have also shown to be directly involved in the epigenetic silencing of transcriptionally silent transgenes in maize (McGinnis et al. 2006). *mop1*, *mop2* and *rmr6* are mutations of known elements within the RdDM pathway (see 1.4.2), knocking out RDR2, NRPD1-like (Pol IV subunit) and NRDP2/E2-like (Pol IV and V subunit) respectively (Arteaga-Vazquez and Chandler 2010). All three affect the accumulation of 24 nt siRNAs *in vivo* and are associated with

loss of DNA methylation. However, genome-wide analysis conducted on *mop1* mutants has shown retrotransposons and DNA TEs to be differentially regulated (Jia et al. 2009) and has shown the removal of DNA methylation following the reversal of silencing (for example of, *Mutator* (*Mu1*) elements) (Lisch et al. 2002), demonstrating a role in regulating TEs and genes. Both *rmr6* and *mop2* (semi-dominant mutation available, *Mop2-1*) are close to *mop1* in RdDM , encoding the Pol IV and Pol V subunits required for the initial steps of the pathway. The study of Pol IV mutants however, has mostly been conducted in *Arabidopsis,* with *nrpd1a* and *nrpd2* mutants (Pol IV and V subunits) showing great loss in DNA methylation in heterochromatin (Onodera et al. 2005). *mop2* is also referred to as *rmr7*, an allele of *mop2* that was under investigation at the same time by different research teams (Stonaker et al. 2009).

The *rmr1* and *rmr2* mutants, identified through genetic screens with the *pl1* allele, however, have less defined roles. The *rmr1* gene is required for the efficient accumulation of 24 nt siRNAs, just as the genes previously described, but encodes an SNF2 protein thought to interact with chromatin marks (Hale et al. 2007). A more recent study also indicates that *rmr1* operates upstream and independently of Pol IV in the RdDM pathway and the loss of which induces hypomethylation of *Mutator* elements (Hale et al. 2009). The role of *rmr2* is less clear but it is known to influence the accumulation of 24 nt siRNAs and the maintenance of DNA methylation at distinct loci, such as the 3' end of *Pl1-Rh* (Barbour et al. 2012). Greater understanding of this element may only arise through the discovery of directly interacting proteins.

In addition, other mutants demonstrate potential roles in epigenetic TGS, with less well understood functions. Firstly, *Unstable factor for orange 1 (Ufo1*), a *trans*-acting modifier of *P1-wr* alleles which controls phlobaphene pigmentation in vegetative and reproductive tissues, was shown to release epigenetic repression and induce hypomethylation (Chopra et al. 2003). Moreover, recently the role of *Ufo1* in paramutation was probed, observing its requirement for the paramutation of *b1* and *p1* loci (Sekhon et al. 2012). Current evidence also suggests that *Ufo1* has roles in multiple epigenetic pathways, due to the loss of H3K9me2 from the *P1-wr* allele following reactivation by this mutation.

morpheus' molecule 1-like (*mom1-like*) is the homologue of MOM1, which, is required for the maintanence of TGS in *Arabidopsis*. Work conducted has shown that the mutation of *mom1* results in loss of methylation from previously silenced genes and heterochromatic repeat regions (Amedeo et al. 2000). Interestingly, changes in expression of genes were also observed, without alterations in DNA methylation, indicating that *mom1* acts independently of the RdDM pathway or downstream (Amedeo et al. 2000; Saze et al. 2012).

Finally, *Rpd3-like* encodes a histone deaceytlase in maize required for gene silencing, similar to a mammalian protein which is recruited by the human retinoblastoma tumour suppressor gene (Rossi et al. 2003). The focus of this study found maize genes that can recruit the plant equivalent, *Rpd3-like*, for the same result. Earlier studies have also found roles for similar genes in *Drosophila* associated with PEV, altering gene expression (De Rubertis et al. 1996). However, little else is known of the role of this protein in plant gene silencing networks.

The RdDM and TGS mutants described have been employed to reactivate NYR-v expression in maize, releasing reporter repression to identify the mechanisms involved.

Table 11: Epigenetic maize mutants affecting siRNA biogenesis and/or TGS

Name	**Abrev.**	**Pathway**	**Function**	**Reference**
required to maintain repression 1	*rmr1*	RdDM	Sucrose nonfermenting 2 (SNF2)-like adenosine triphosphatase chromatin remodeller	Hale et al., 2007
required to maintain repression 2	*rmr2*	unknown	unknown	Barbour et al., 2012
required to maintain repression 6	*rmr6*	RdDM	NRPD1-like (largest subunit of DNA-dependent RNA polymerase IV and V)	Erhard et al., 2009
mediator of paramutation 1	*mop1*	RdDM	RNA-dependent RNA polymerase 2 (RDR2)	Dorweiler et al., 2000
mediator of paramutation 2	*mop2*	RdDM	NRPD2/E2-like (second largest sububit of DNA-dependent RNA polymerase IV)	Sideorenko et al., 2009
unstable factor for orange 1	*ufo1*	TGS	unknown	Chopra et al., 2003
morpheus' molecule 1-like	*mom1-like*	TGS	unknown	pers. comm. Biogemma
Rpd-3 like	*Rpd-3 like*	TGS(PEV)	unknown	De Rubertis et al., 1996 Rossi et al., 2003

4.2 Results

4.2.1 Small RNA profiling by next generation sequencing

In order to determine if small RNA induced epigenetic silencing was responsible for silencing of promoter function of NYR-v, SBS deep sequencing was conducted. This allowed for sequencing of all small RNA species (<200 nt) from NYR-v transgenic and WT control plants. The sequencing data generated were then aligned against the complete NYR-v sequence to display small RNAs matching regions of the transgene.

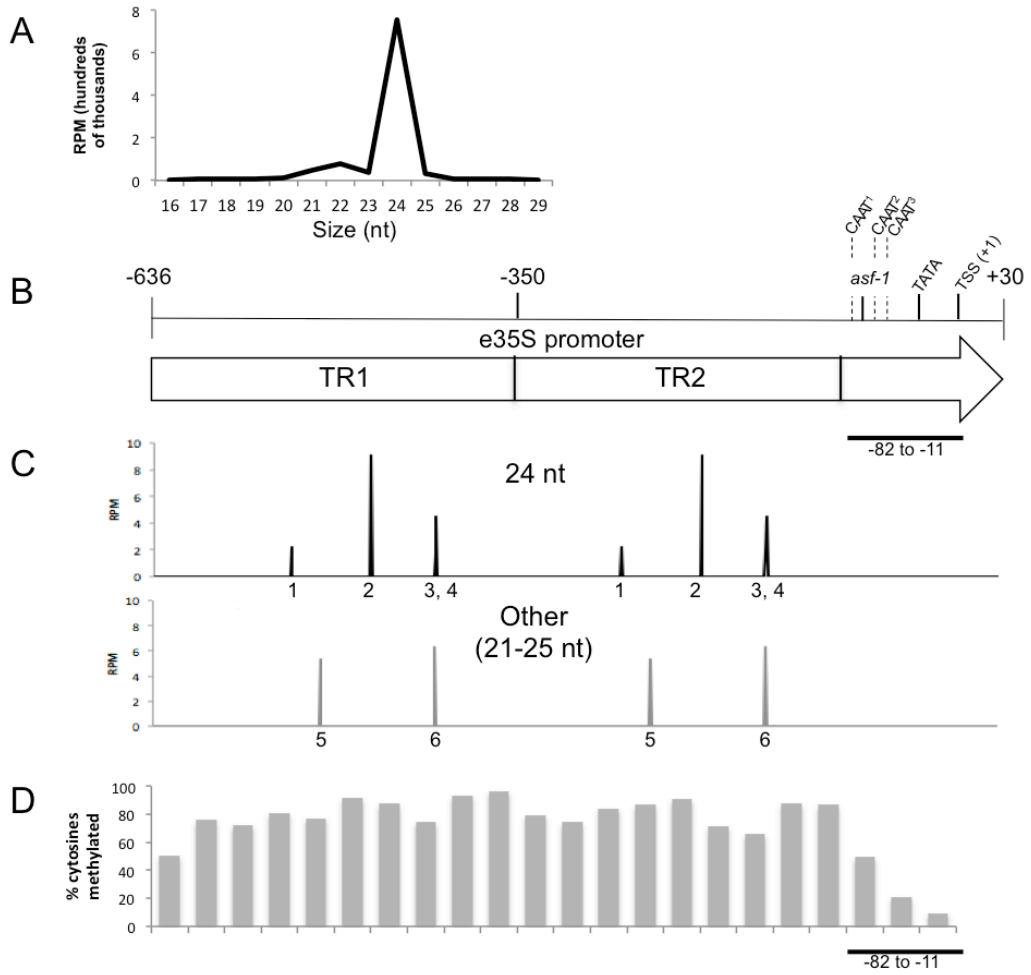
The NYR-v and WT libraries generated comprised of 10,494,510 and 9,447,147 (derived from the same sequencing file of 29,052,106 reads (two samples multiplexed)) reads respectively (Appendix Table.22 and 23). Dissection of small RNA size class revealed 44,333 (4.4%), 75,455 (7.5%), 34,837 (3.5%) , 719,726 (72%) and 32,121 (3.2%) RPM of 21, 22, 23, 24 and 25 nt respectively in the NYR-v data set (Fig.28A, WT control, Appendix Fig.73). Subsequent alignment of the libraries to the NYR-v sequence shows 21-25 nt small RNAs mapping exclusively to the tandem elements of the promoter region (Fig.28C). This includes six sequences recongised in both tandem elements, four of which are 24 nt in size (2.18, 9.10, 1.68 and 4.45 RPM (totaling 17.41 RPM)), one of 21 nt (5.24 RPM) and one of 23 nt (5.93 RPM). The location of small RNAs in the tandem elements of the promoter correlate with the presence of DNA methylation previously identified (Fig.28D).

Attempts to correlate individual cytosine methylation positions with proximity of small RNA alignments proved futile as methylation was of similar density across the whole region (minus the 3' unique distal region) (Fig.28D). The precise position placement of DNA methylation through the action of small RNAs is also unpredictable and not well understood. Small RNAs were also detected, in lower abundance, at other positions in the transgene (Appendix Table.24). However, these sites lack the presence and abundance of the 24 nt size class.

This analysis shows the presence of small RNAs known to be involved in RdDM aligning to the NYR-v promoter.

4.2.2 Introgression into RdDM mutants

4.2.2.1 Preliminary analysis In order to test the hypothesis that small RNAs are involved in NYR-v silencing, a range of RdDM mutants and epigenetic mutants were utilised. Five mutants involved in the RdDM pathway (*mop1-1*, *rmr1-1*, *Mop2-1* (semi-dominant), *rmr6-1* and *rmr2-1*) and three mutants hypothesized to be involved in TGS (*Ufo1-1*, *mom1-like*, *rpd3-like*) were genetically introgressed with NYR-v.



E

No.	Sequence	Strand	Size (nt)	Position	RPM
1	AAGATAGTGGAAAAGGAAGGTGGC	+	24	-527 to -504, -265 to -242	2.18
2	GTTGAAGATGCCTCTGCCGACAGT	-	24	-464 to -441, -202 to -179	9.10
3	AGGAGCATCGTGGAAAAAGAAGAC	+	24	-413 to -390, -151 to -128	1.68
4	GGAGCATCGTGGAAAAAGAAGACG	+	24	-412 to -389, -150 to -127	4.45
5	TCCTACAAATGCCATCATTGC	-	21	-503 to -483, -241 to -221	5.34
6	GGAGCATCGTGGAAAAAGAAGAC	+	23	-412 to -390, -150 to -128	5.93

Figure 28: Small RNA profile of NYR-v performed by next generation sequencing. Sequencing of NYR-v RNA enriched for small RNA classes (>200 nt), (A) Total abundance of small RNAs detected in the library - RPM, Reads per Million. (B) Schematic diagram of NYR-v promoter region - TR, tandem repeat; transcriptional start site (TSS) +1. (C) Abundance and position of 24 nt size class and other sizes (21-25 nt) detected at the promoter region. A number of between 1-6 is assigned to each sequence aligned. (D). Level of DNA methylation present in the promoter region, each bar represents the average percentage of total cytosines methylated encompassing seven cytosine positions. Black line denotes region of reduced methylation mirrored in NYR-v schematic (B). (E) Small RNA sequence, strand alignment, size, position and abundance. RPM, Reads per Million.

RdDM Mutant introgression was carried out as displayed by the crossing strategy in Fig.29. Transgenics hemizygous for NYR-v were crossed with homozygous mutants generating NYR heterozygous mutant I_1 progeny, this was in turn sibling crossed generating I_2 progeny. Homozygous mutant NYR plants were then crossed by heterozygous mutant NYR over two subsequent generations, producing I_3 and I_4 progenies. Crossing of homozygous and heterozygous mutants was performed in this manner due to the problems associated with seed abortion and seed size if both parents are homozygous for *mop1-1* or *rmr1-1*. Furthermore, plant growth and crosses for mutant introgressions were performed away from facilities possible to conduct NYR reporter expression analysis, therefore NYR expression was unknown between generations.

Backcrossing of I_4 homozygous mutant NYR by WT generated BC_1 progeny. BC_1 Heterozygous mutant NYR was crossed with WT again producing BC_2 progeny, which was in turn backcrossed a third time utilsing BC_2 non-mutant NYR generating BC_3 progeny. The crosses involved in generating the BC progenies will be discussed in greater detail at a later stage.

Initial screening of reporter expression was conducted by confocal microscopy with I_1 and I_2 transgenics, analysing the roots of newly germinated seeds revealing reactivation (Fig.30). NYR-v was observed alongside NYR-a* (NYR-a* represent reactivated NYR-v plant) in progenies from all mutant introgressions in I_2 (Fig.31). The F_1 WT background showed 0.0% reactivated, a result shared by all I_1 RdDM mutants tested, except for I_1 *Mop2-1* which showed 14.5% of progeny reactivated. Additionally, the I_2 *mop1-1* showed 8.3% and 19.8% in homozygous and heterozgygous respectively, I_2 *rmr1-1* showed 50.0% and 40.4% in homozygous and heterozygous respectively, I_2 *Mop2-1* showed 21.3% and 23.0% in homozygous and heterozygous respectively, I_2 *rmr6-1* showed 7.1% and 11.1% in homozygous and heterozygous respectively, I_2 *rmr2-1* showed 64.4% and all others were indistinguishable from WT. Introgression of NYR-v with *Ufo1-1*, *mom1-like* and *Rpd3-like* resulted in no progeny plants showing reactivation.

These data showed that introgression with RdDM mutants allowed for reactivation of the NYR-v transgene in a proportion of progeny plants.

4.2.2.2 Further mutant introgression Following from NYR-v reactivation in RdDM mutant backgrounds introgression of a further two generations was conducted (I_3 and I_4) (Fig.29) in an attempt to increase the proportion of reactivated progeny. This has been shown in previous work conducted, whereby percentage of plants showing the release of silencing of *mudrA* increased over multiple generations in *mop1-1* mutant background (Woodhouse et al. 2006). This has been hypothesised to reduce epigenetic modifications, increasing the abundance of decayed epigenetic states (Sekhon and Chopra 2009). As before, confocal microscopic analysis of roots was carried

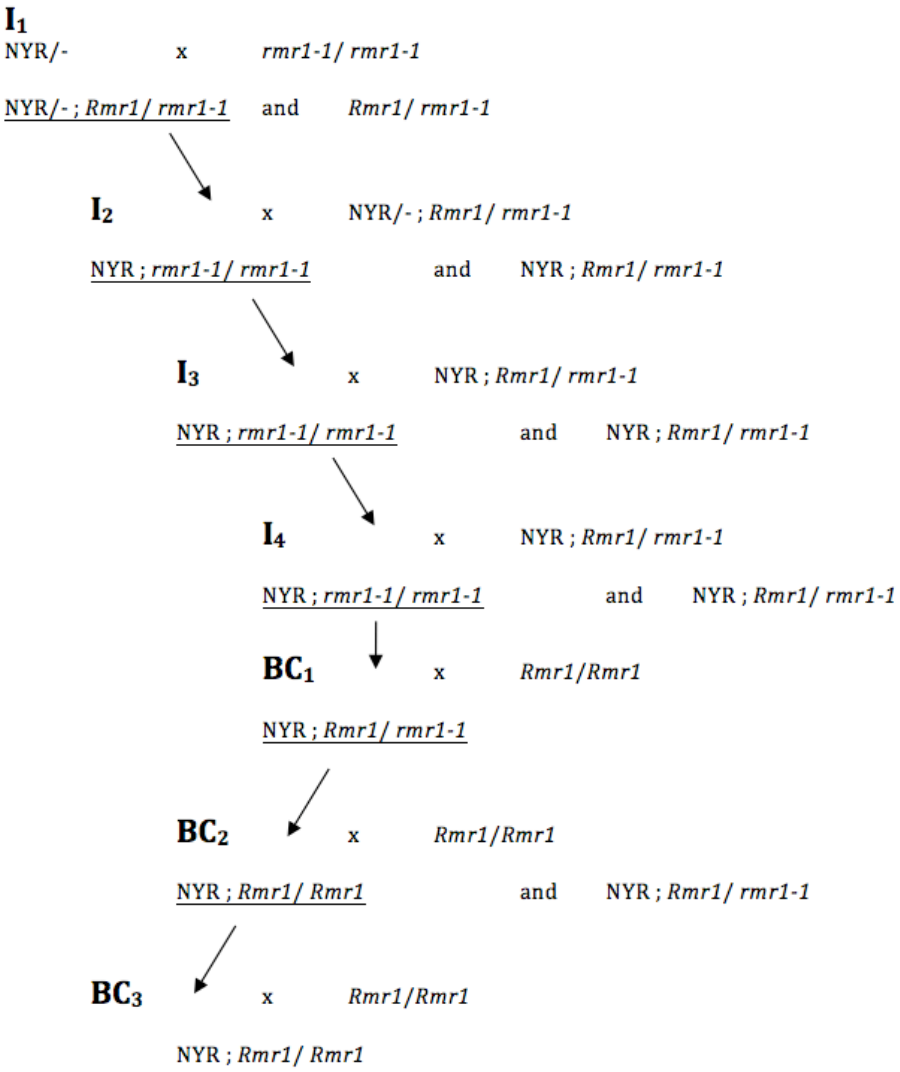


Figure 29: Crossing strategy for mutant introgression into NYR-v plants. NYR-v plants were crossed with homozygous RdDM mutants (example: *rmr1-1* introgression) and subsequently crossed with mutants for up to four generations (I_1 to 1_4). Following the fourth introgression NYR was outcrossed with B73 WT over three generations generating BC_1 , BC_2 and BC_3 progenies. Zygosity of NYR transgene unknown after initial cross, genotyping based on presence or absence of transgene. I, Introgression generation; BC, Backcross generation.

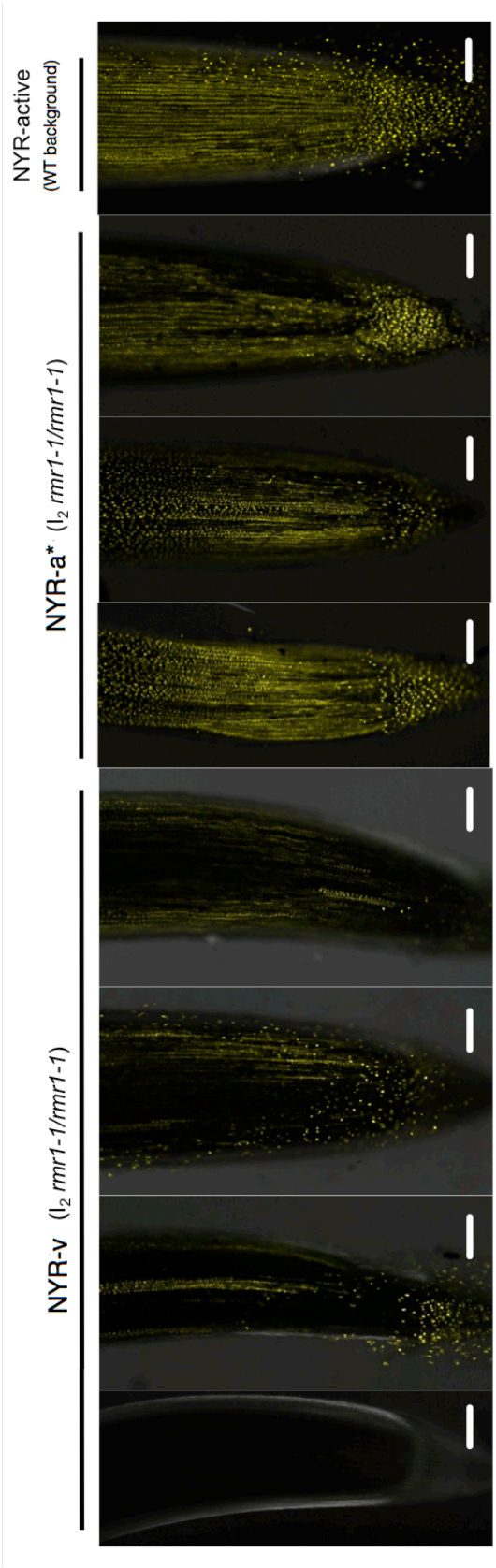


Figure 30: Confocal microscopic analysis of I_2 *rmr1-1/rmr1-1* mutant progeny and WT roots. From left to right, a single NYR-s, three examples of NYR-v, three examples of NYR-a* reactivated by *rmr1-1* and a single NYR-a from a WT line. Scale bar = 200 µm

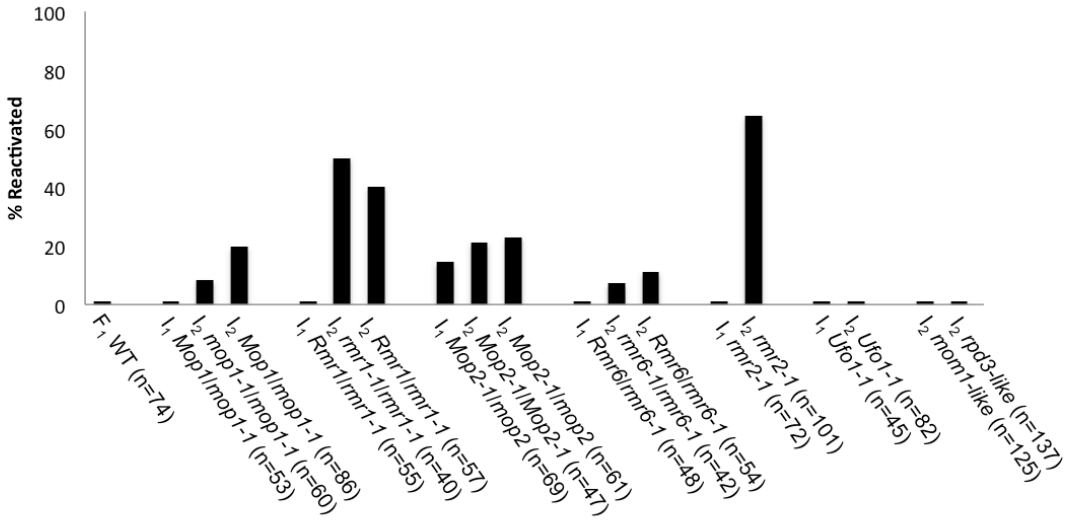


Figure 31: RdDM mutant introgression screen showing percentage of NYR-a* reactivated progeny as detected in the roots/leaves of different RdDM mutants. Analysis of roots from newly germinated seeds, classifying as active or variegated in WT and I_2 mutant/RNAi backgrounds dependent on YFP expression potency/location.

out, collecting data on I_3 and I_4 *mop1-1*, *rmr1-1* and *Mop2-1* backgrounds (Fig.32). The WT background demonstrates 0.0% reactivation over F_1, F_2 and F_3 progenies indicating that variegated expression is heritable and stable. The *mop1-1* background shows 19.8%, 27.5% and 23.8% in heterozygotes and 8.3%, 27.5% and 51.7% in homozygote over I_2, I_3 and I_4 progenies. The *rmr1-1* background shows 40.4%, 73.3% and 80.0% in heterozygotes and 50.0%, 82.5% and 68.4% in homozygote over I_2, I_3 and I_4 progenies. The *Mop2-1* background shows 23.0%, 26.3% and 23.9% in heterozygotes and 21.3%, 30.2% and 29.2% in homozygote over I_2, I_3 and I_4 progenies.

Collectively these data indicates that introgression into the RdDM mutant backgrounds, *mop1-1* and *rmr1-1,* over multiple generations increases the proportion of plants with reactivated NYR-v.

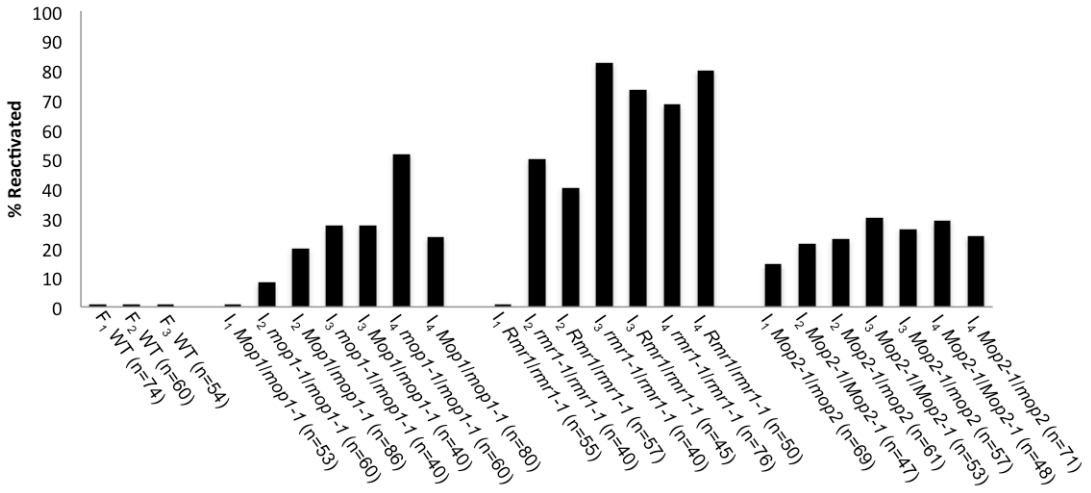


Figure 32: Percentage of NYR-a* progeny identified as reactivated in the roots/leaf following multiple generations of RdDM mutant introgression. Analysis of roots from newly germinated seeds, classifying as active or variegated in WT and mutant backgrounds dependent on YFP expression potency/location over three generations. WT (F_1 – F_3) and *mop1-1*, *rmr1-1* and *Mop2-1* mutants (I_2 – I_4).

4.2.3 Transcriptional analysis of NYR in different RdDM mutant backgrounds

To confirm a release from silencing in different RdDM mutant backgrounds, I_4 (and *rmr6-1* I_2) transgenics were analysed by qRT-PCR, measuring expression of YFP (Fig.33). In addition, I hypothesised that loss of RdDM function would relax chromatin in flanking sequences resulting in transcription. This was assessed through detection of potential LTR UPN (L-UPN) readthrough transcripts, utilising priming sites at the beginning and end of the non-coding UPN region. The detection of which had previously been shown by PCR and gel electrophoresis (data not shown).

This analysis revealed that NYR-a* homozygous for *mop1-1*, *rmr1-1*, *Mop2-1* and *rmr6-1* mutants showed 2.4, 3.3, 3.6 and 2.3 $\log_2$ fold change in expression of YFP respectively compared with NYR-v heterozygous *mop1-1*, *rmr1-1*, *Mop2-1* and *rmr6-1* mutants, which showed 0.2, 1.1, 0.7 and -0.2 $\log_2$ fold change respectively. Additionally, NYR-a* homozygous *mop1-1*, *rmr1-1*, *Mop2-1* and *rmr6-1* mutants showed 6.7, 5.8, 7.6 and 5.9 $\log_2$ fold change respectively in detected expression of the L-UPN readthrough transcript, compared with NYR-v heterozygous *mop1-1*, *rmr1-1*, *Mop2-1* and *rmr6-1* mutants showing -1.7, -0.3, 0.9 and 1.27 $\log_2$ fold change respectively.

These results confirm the increased mRNA abundance of the proportion of NYR-a* plants produced, following introgression into RdDM mutant backgrounds, suggesting that NYR-v is regulated epigenetically.

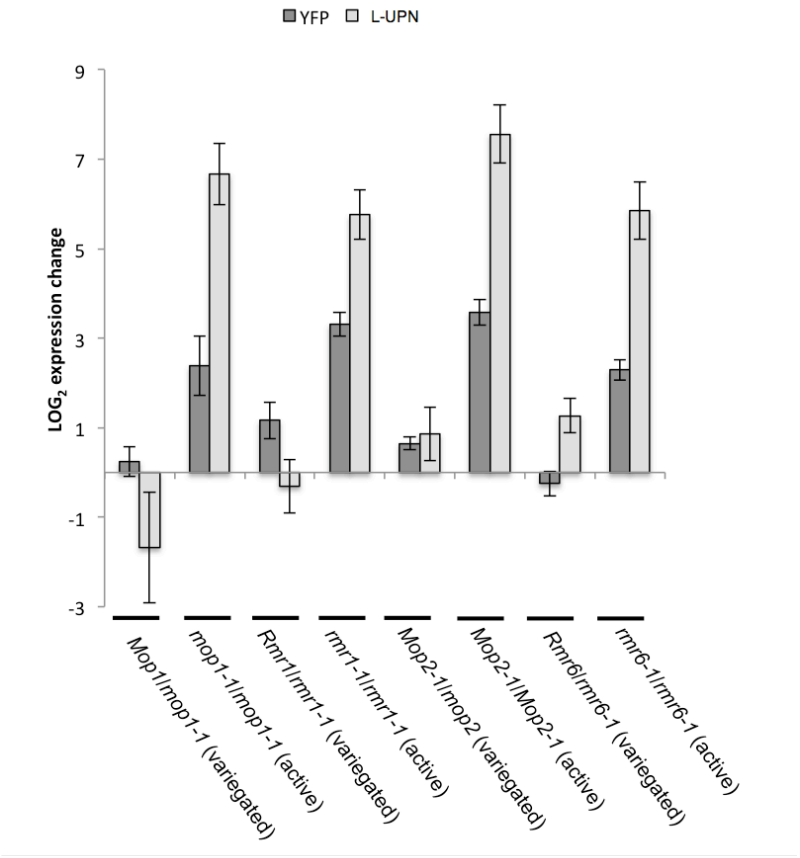


Figure 33: qRT-PCR expression analysis of YFP and L-UPN readthrough transcript in different RdDM mutant backgrounds. Expression of homozygous (active) and heterozygous (silenced) mutant backgrounds, *mop1-1* (I_4), *rmr1-1* (I_4), *Mop2-1* (I_4) and *rmr6-1* (I_2) compared against NYR-v WT, RNA derived from leaf 4 tissue. Differential expression is shown as $\log_2$ fold change.

4.2.4 Methylation analysis of NYR in different RdDM mutant backgrounds

4.2.4.1 Methylation analysis by McrPCR DNA methylation analysis was conducted in three regions of NYR in WT and I_4 (and *rmr6-1* I_2) RdDM mutant backgrounds (Fig.34). This allowed the analysis of methylation at the region immediately upstream of the promoter and the functional elements of the transgene, namely the e35S promoter and the YFP reporter. NYR-v in the WT background shows hypomethylation in the 5' region, hypermethylation in the e35S promoter region and hypomethylation in the YFP region. All NYR-v heterozygous RdDM mutants (*mop1-1*, *rmr1-1*, *Mop2-1* and *rmr6-1*) present the same profile as the WT background, whereas NYR-a* homozygous RdDM mutants (*mop1-1*, *rmr1-1*, *Mop2-1* and *rmr6-1*) differ in showing hypomethylation in the e35S promoter region (Fig.34).

However, it is important to recognise that this analysis does not investigate methylation of NYR-a* in heterozygous mutants or NYR-v plants derived from homozygous mutants. Therefore, DNA methylation analysis was also conducted on the e35S promoter of NYR-v in the WT background and NYR-v and NYR-a* in homozygous RdDM mutants (*mop1-1*, *rmr1-1*, and *Mop2-1*) over three generations (F_1 to F_3 (WT), I_2 to I_4 (RdDM mutants)). The analysis was also conducted on NYR-v and NYR-a* in homozygous *rmr6-1* and *rmr2-1* (unknown genotype) for a single generation (I_2). All NYR-v in the WT background show hypermethylation of the e35S promoter across all three generations and all NYR-v and NYR-a* in RdDM mutant backgrounds show hypomethylation, regardless of NYR expression, except *rmr2-1* (Fig.35 and 36). *rmr2-1* shows hypomethylation in some of the NYR-v which may be due to a mix of homozygous and heterozygous individuals utilised.

These data show that there is no direct correlation, that can be detected by McrPCR analysis, for reactivation of NYR-v and the DNA methylation at the whole of the promoter region.

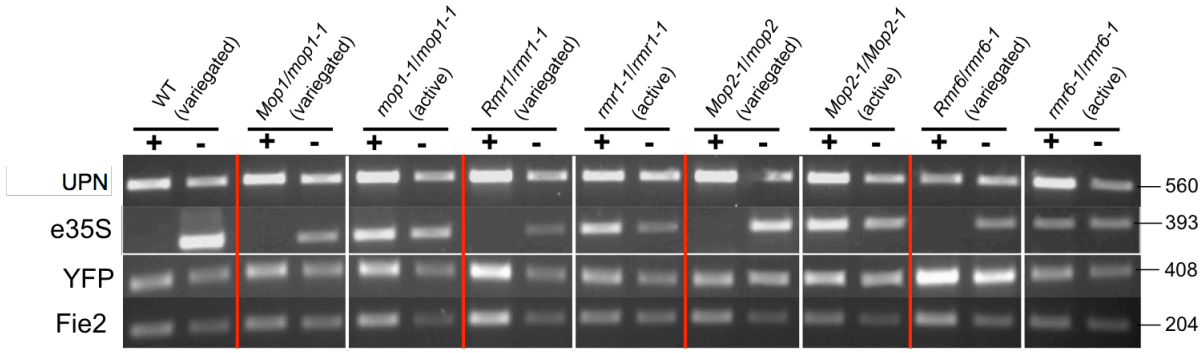


Figure 34: McrPCR methylation analysis of NYR in different RdDM mutant backgrounds. DNA methylation analysis of UPN region, e35S promoter, YFP and *Fie2* (unmethylated control) in WT, *mop1-1* (I_4), *rmr1-1* (I_4), *Mop2-1* (I_4) and *rmr6-1* (I_2) heterozygous (variegated) and homozygous (active) mutants. gDNA derived from leaf 4 tissue. PCR product sizes (in bp) is indicated on right hand side of the panel.

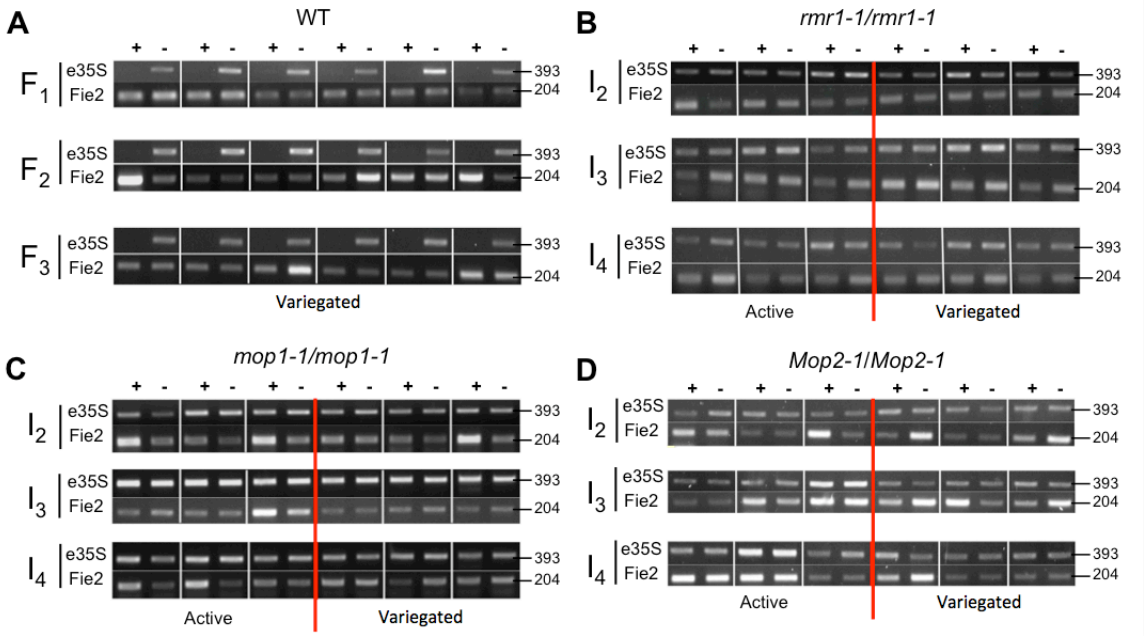


Figure 35: McrPCR methylation analysis of the NYR promoter region over three generations of RdDM mutant introgression. McrBC digestion and PCR amplification of the NYR e35S promoter region and Fie2 (unmethylated control in (A) six NYR-v WT plants from each of three generations (F_1-F_3), (B-D) three NYR-v and three NYR-a* in homozygous mutant background across three generations of introgression (I_2-I_4). Mutants are *rmr1-1*, *mop1-1* and *Mop2-1*. PCR product sizes (in bp) is indicated on right hand side of the panel.

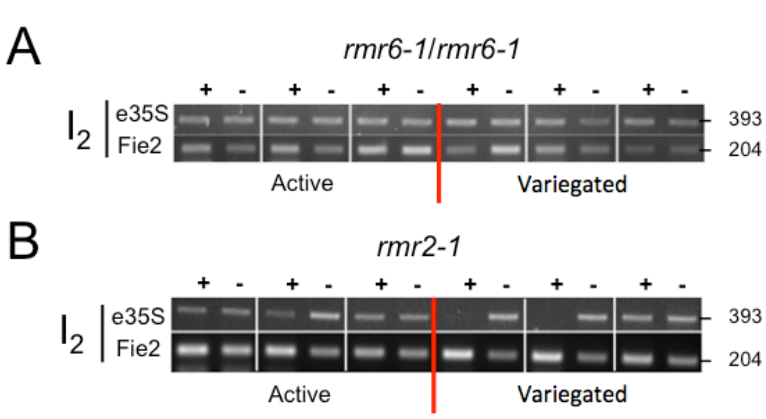


Figure 36: McrPCR methylation analysis of NYR promoter region in additional RdDM mutants. McrBC digestion and PCR amplification of the NYR e35S promoter and Fie2 (unmethylated control) in three NYR-v and three NYR-a* plants of an I_2 RdDM mutant background. (A) *rmr6-1* (homozygous) and (B) *rmr2-1* (unknown zygosity). PCR product sizes (in bp) is indicated on right hand side of the panel.

4.2.4.2 Methylation analysis by bisulfite sequencing To obtain a detailed map of DNA methylation at the NYR transgene in RdDM mutants, bisulfite sequencing was utilised targeting the promoter region. McrPCR has already indicated that both NYR-a* and NYR-v homozygous RdDM mutants showed hypomethylation of the promoter, therefore, specific sites of methylation may be important in silencing. Bisulfite sequencing was conducted on NYR-a* homozygous RdDM mutants, compared with NYR-v in a WT background.

The bisulfite sequencing data reinforces the data obtained in the McrPCR analysis showing a significant reduction in methylation across the promoter region in NYR-a* homozygous *mop1-1*, *rmr1-1*, *Mop2-1* and *rmr6-1* mutants which show 26.6%, 32.7%, 31.9% and 37.9% average methylation respectively, compared with the NYR-v WT showing 73.0% methylation (Fig.37). The reduction is noticeable at all CG, CHG and CHH methylation sequence contexts.

When the methylation level was analysed for individual cytosine locations across the promoter region, a uniform reduction can be observed in the NYR-a* RdDM mutants compared with NYR-v in wild type background (Fig.38). Two regions were identified with a greater reduction in methylation to neighbouring cytosine positions, at -488 to -408 and -176 to -156. Both regions exist within the tandem elements and share sequence overlap, therefore reduction of methylation is affecting both areas similarly. Both regions are also in close proximity to 24 nt small RNAs identified previously (Fig.28B).

The 3' distal unique region (positions -87 to +1) also show a reduction in methylation of the NYR-a* RdDM mutants (Fig.38). The region, comprising of 24 cytosine positions, is on average

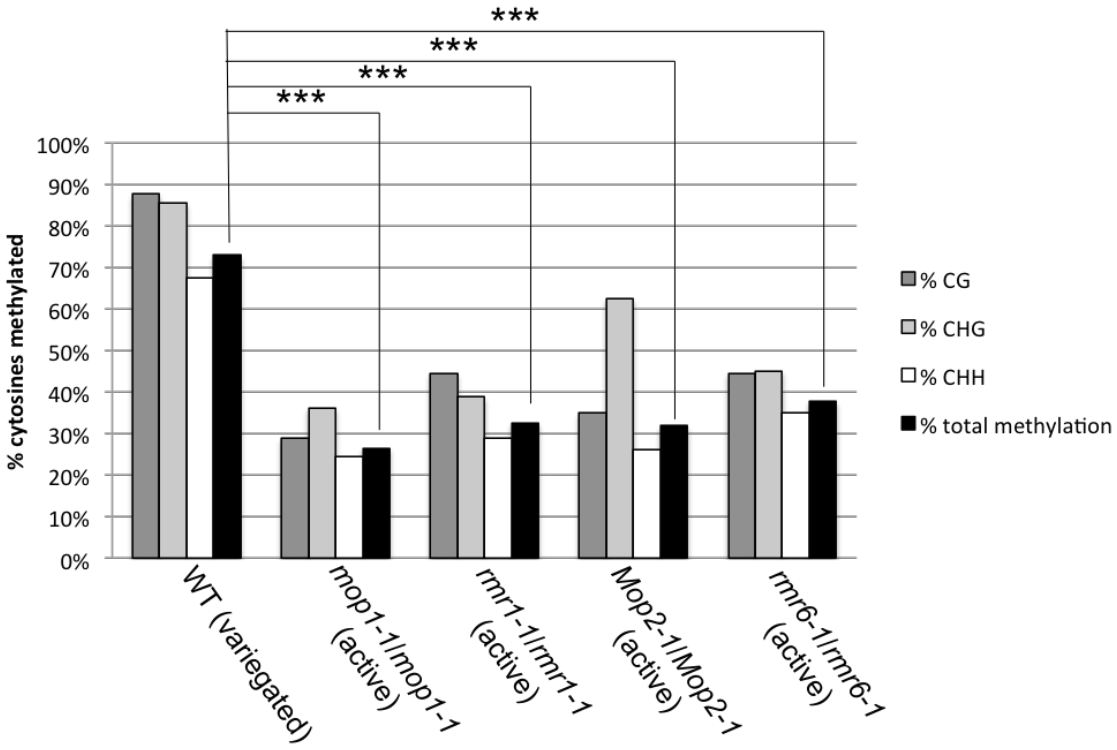


Figure 37: Bisulfite sequencing analysis of the NYR promoter in WT and RdDM mutant backgrounds. PCR amplification and sequencing of bisulfite treated gDNA focusing on the NYR e35S promoter region of WT (NYR-v), I_4 *mop1-1/mop1-1*, I_4 *rmr1-1/rmr1-1*, I_4 *Mop2-1/Mop2-1* and I_2 *rmr6-1/rmr6-1* (NYR-a*) plants. The level of methylation is reported as the average percentage of total cytosines exhibiting methylation established from 15 independent clones for each data set from position -635 to -11 of NYR transgene, a total of 154 cytosines. Asterisks denote significant difference between silenced WT and active mutant levels of methylation ($p<0.001$) as established by Wilcoxon two sample rank test. Raw data presented in Appendix Table.25 and Lollipop diagram of methylation state of individual cytosine positions in all clones – Appendix Fig.74 and 75. ***, $p<0.001$.

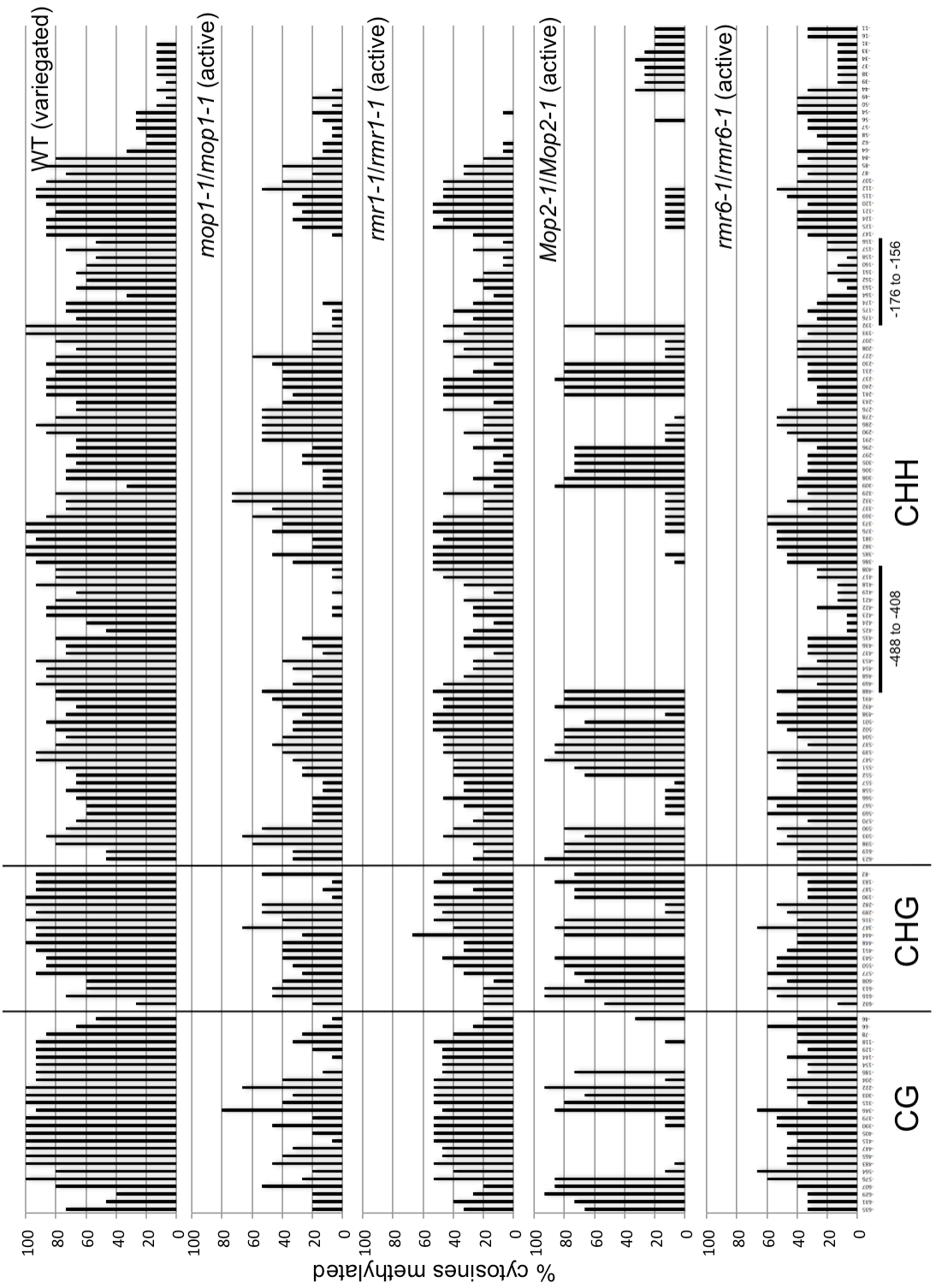


Figure 38: Methylation analysis of NYR e35S promoter showing CG, CHG and CHH methylation in WT and RdDM mutant backgrounds. Bisulfite analysis of DNA methylation in the NYR e35S promoter in WT (NYR-v), I_4 *mop1-1/mop1-1*, I_4 *rmr1-1/rmr1-1*, I_4 *Mop2-1/Mop2-1* and I_2 *rmr6-1/rmr6-1* (NYR-a*) plants. The level of methylation is reported as the percentage of individual cytosine locations exhibiting methylation, established from 15 independent clones, spanning -635 to -11 of NYR (cytosine positions marked on the x-axis and have been split between CG, CHG and CHH contexts), a total of 154 cytosines. Positions -488 to -408 and -176 to -156 (underlined) represent regions of conserved reduction in methylation between mutants.

33.3% methylated in the WT background and significantly reduced in NYR-a* homozygous *mop1-1*, *rmr1-1* and *Mop2-1* and *rmr6-1* mutant backgrounds showing 11.9%, 10.0%, 15.0% and 30.8% methylated respectively. In addition, methylation of two individual CG positions within the *asf-1* region (Kanazawa et al. 2007a) (positions -78 and -66) show 86.7% and 66.7% respectively in the WT background and 26.7% and 13.3%, 40.0% and 26.7%, 0.0% and 0.0%, 40.0% and 60.0% methylation in the *mop1-1*, *rmr1-1* and *Mop2-1* and *rmr6-1* mutants respectively. Analysis of all cytosine positions in the *asf-1* region show an average of 70.0% methylation in the WT background and 26.7%, 30.0%, 18.3% and 45.0% methylation in *mop1-1*, *rmr1-1* and *Mop2-1* and *rmr6-1* mutants respectively. The CAAT[1]-like element which shares overlap with the *asf-1* region also shows a reduction, averaging 86.7% methylation in the WT and 37.8%, 33.3%, 24.4% and 37.8% in the *mop1-1*, *rmr1-1* and *Mop2-1* and *rmr6-1* mutants respectively.

As before the methylation profile of individual clones were observed across the whole NYR e35S promoter and just at the *asf-1* element (Fig.40). These data show a reduction in the methylation of individual clones between NYR-v in a WT background and the NYR-a* RdDM homozygous mutants across the whole promoter region as expected due to the lower average methlyation observed. The *asf-1* element also shows this reduction with a greater number of clones exhibiting no methylation or methylation of a single cytosine compared with NYR-v in the WT background.

The loss of DNA methylation at the NYR-a* promoter in homozygous RdDM mutants compared with NYR-v in a WT background indicates that these genes are required in part for the maintenance of silencing, however, the extent of methylations involvement remains unclear.

4.2.5 Restoration of RdDM function

After introgression of NYR into RdDM mutant backgrounds, crosses with B73 wild type plants were performed to restore RdDM function. This allowed for the study of the importance of the different RdDM components in the reestablishment of NYR silencing. Previous studies have often shown that RdDM mutants can reactivate silenced transgenes and that reactivation can be reversed upon restoration of RdDM components (McGinnis et al. 2006). The crossing strategy used is represented in Fig.41, the initial cross (BC_1) utilising an I_4 NYR-a* homozygous RdDM mutant and B73 WT resulted in heterozygous mutant progeny (Fig.42). The *mop1-1* background shows 51.7% and 23.8% of progeny reactivated in I_4 homozygous and heterozygous mutants respectively and 11.9% in BC_1(heterozygous). The *rmr1-1* background shows 68.4% and 80.0% of progeny reactivated in I_4 homozygous and heterozygous mutants and 81.8% in BC_1(heterozygous). The *Mop2-1* background shows 29.2% and 23.9% of progeny reactivated in I_4 homozygous and heterozygous mu-

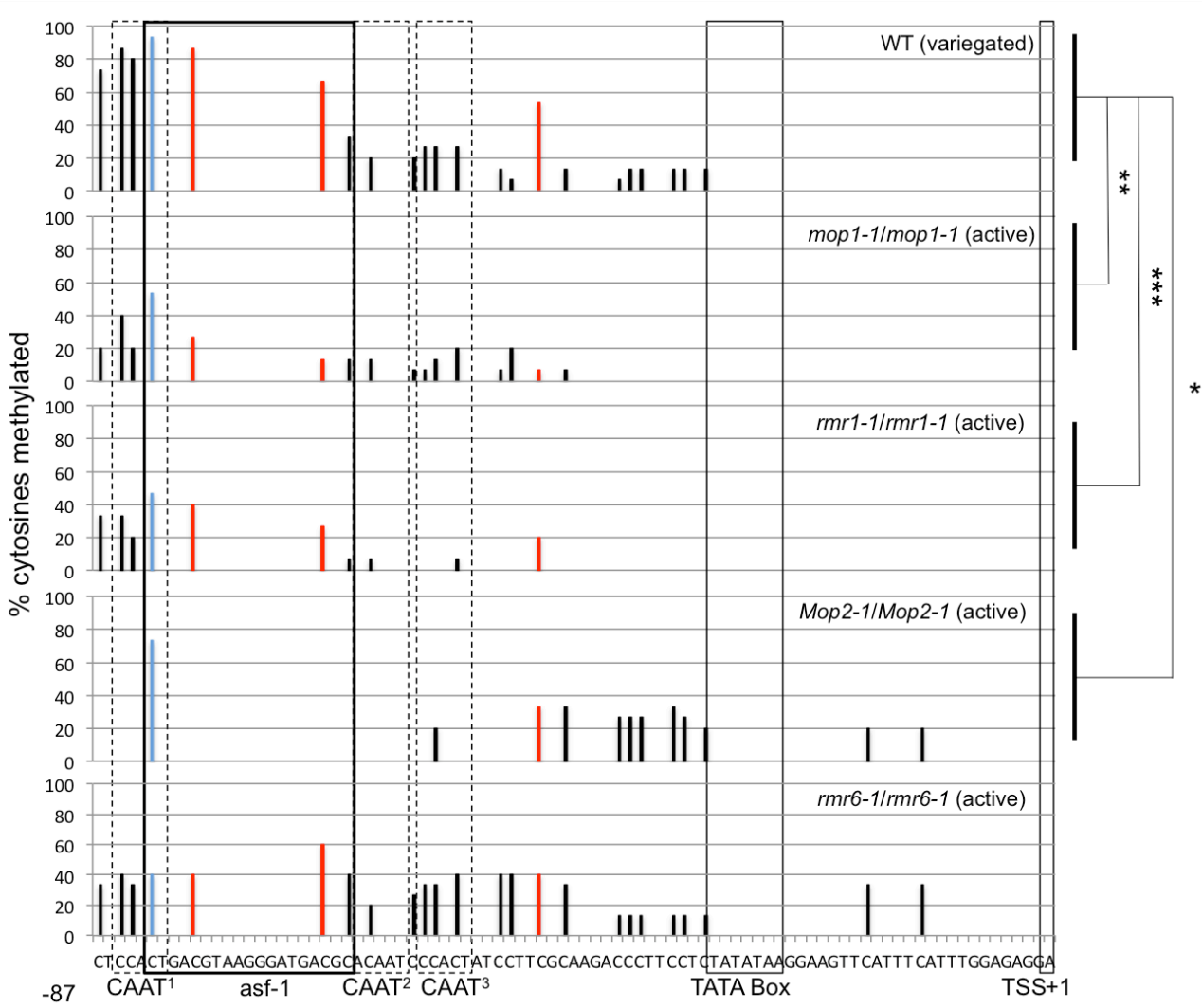


Figure 39: DNA Methylation profile of NYR 3' promoter region in WT and RdDM mutants. Methylation levels of cytosine positions between -87 and +1 of the NYR e35S promoter, displaying CAAT-like elements, *asf-1* element, TATA box and TSS (+1) in WT (variegated) and homozygous RdDM mutant (active) backgrounds. Red representing CG, blue as CHG and black as CHH methylation. The level of methylation is reported as the percentage of individual cytosine locations exhibiting methylation, established from 15 independent clones (position number marked on the x-axis). The asterisks denote the significant difference between average levels of methylation ($p<0.05$) of the data sets as established by Wilcoxon two sample rank test. *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$.

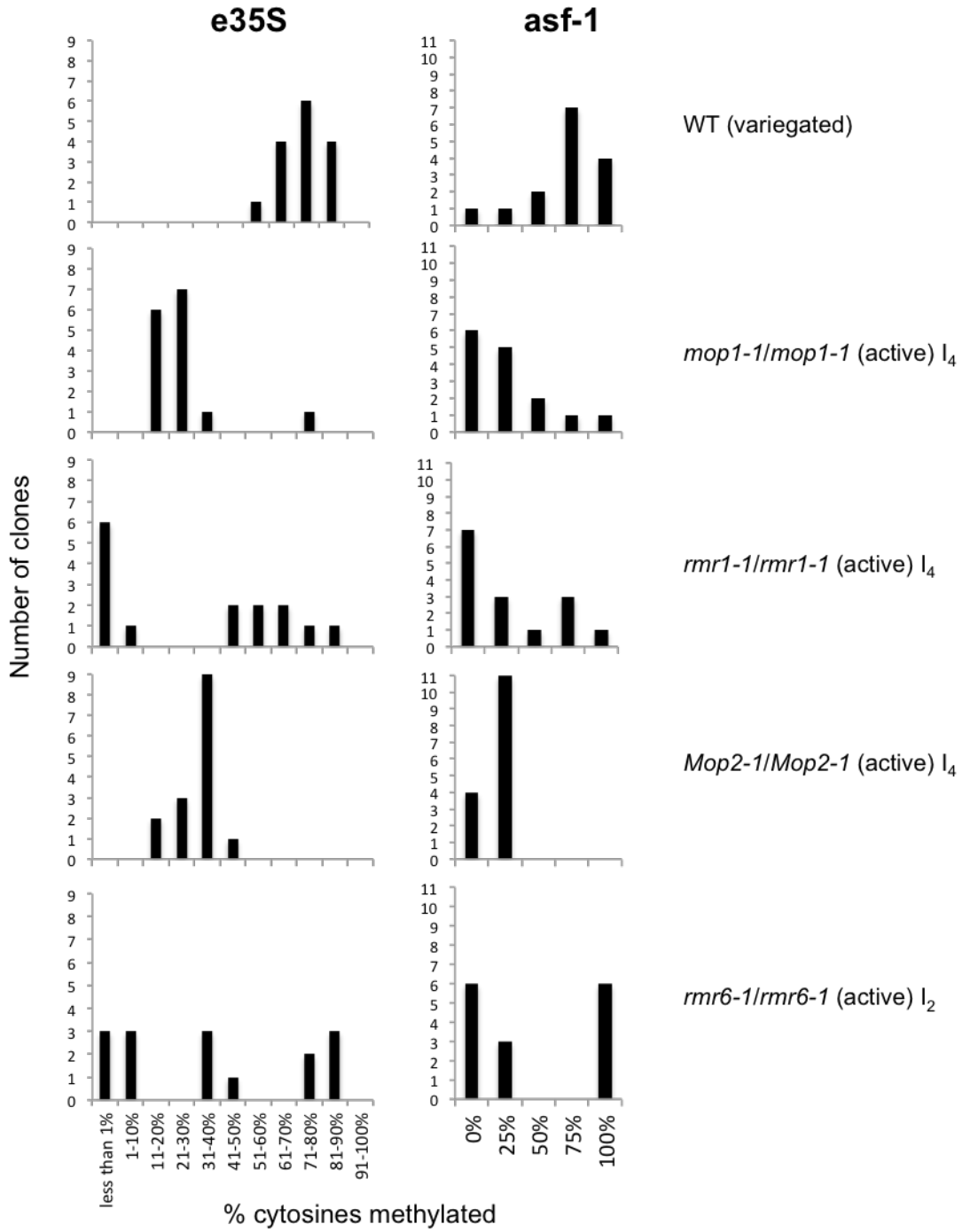


Figure 40: DNA methylation status of individual clones across the NYR e35S promoter in WT and RdDM mutant backgrounds. The level of methylation for individual clones is reported as a percentage of cytosines exhibiting methylation in the NYR e35S promoter and the asf-1 region of the promoter (only four cytosines) in WT (NYR-v), I_4 *mop1-1/mop1-1*, I_4 *rmr1-1/rmr1-1*, I_4 *Mop2-1/Mop2-1* and I_2 *rmr6-1/rmr6-1* (NYR-a*) plants. 15 clones had been sequenced for each dataset.

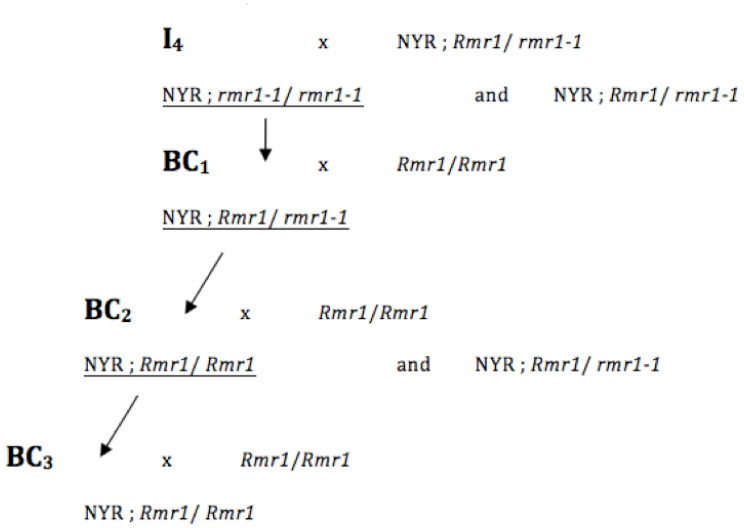


Figure 41: Crossing strategy for backcrossing NYR plants previously introgressed with RdDM mutants. NYR-a* plants homozygous for RdDM mutants derived from the I_4 generation were crossed with B73 WT plants over three generations, generating BC_1, BC_2 and BC_3 progenies. At each cross an NYR-a* plant was selected for crossing. The BC_1 generation was conducted for *mop1-1*, *Mop2-1* and *rmr6-1* (I_2 in the case of *rmr6-1*), while *rmr1-1* (shown as an example in the figure) was taken until the BC_3 generation. I, Introgression; BC, Backcross generation.

tants and 21.7% in BC_1(heterozygous). Finally, the *rmr6-1* background shows 7.1% and 11.1% of progeny reactivated in I_2 homozygous and heterozygous mutants and 4.5% in BC_1(heterozygous). These data show there was no reversion to all NYR plants showing variegated expression in BC_1 progenies, NYR-a* plants remained. This result was expected as previous analysis has shown that NYR-a* expressing plants exist in heterozygous RdDM backgrounds. However, it was unknown if crossing outside of mutant backgrounds altogether would produce a different outcome.

A second round of backcrossing was performed utilising BC_1 NYR-a* heterozygous for *rmr1-1* generating BC_2 progenies. The heterozygous *rmr1-1* NYR-a* BC_1 progeny was utilised, as it was thought that it could show the most noticeable drop in proportion of NYR-a* expressing plants if silencing was to return, as it possessed the highest proportion (81.8%) (Fig.43A). The resulting heterozygous *rmr1-1* and non-mutant progeny of BC_2 show 67.5% and 41.9% reactivation respectively. A non-mutant NYR-a* plant from BC_2 was then backcrossed again generating the BC_3 progeny, which showed 100.0% reactivation.

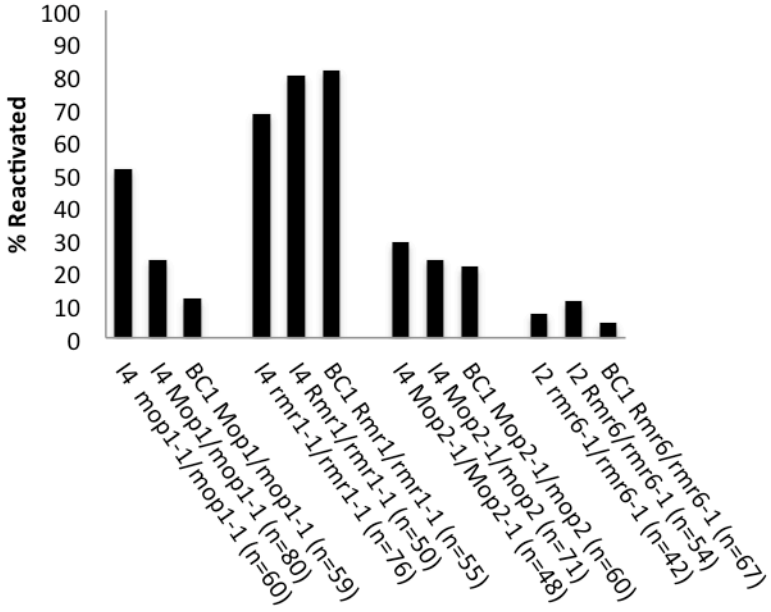


Figure 42: NYR-a* progeny identified in BC_1 progenies Percentage of progeny reactivated, detected through confocal microscopic analysis of roots following cross of I_4 (I_2 in the case of *rmr6-1*) NYR-a* in a homozygous mutant background with B73 WT, generating BC_1 populations.

Furthermore, to assess if methylation returned with the restoration of *rmr1* I conducted McrPCR analysis (Fig.43B). At BC_1, three heterozygous *rmr1-1* NYR-a* and three heterozygous *rmr1-1* NYR-v individuals were tested, a single NYR-a* plant showed hypomethylation in the promoter region and five others (both NYR-a* and NYR-v) showed hypermethylation. At BC_2, eight non-mutant NYR-a* plants were tested, showing two as hypomethylated and six as hypermethylated. Finally, the BC_3 progeny, which tested non-mutant NYR-a* progeny, seven individuals were hypomethylated and only a single was hypermethylated.

These data indicate that the NYR-a* epigenetic state derived from *rmr1-1* mutation is heritable over multiple generations despite restoration of *rmr1* function. Furthermore, the NYR-a* expressing plants generated show both hypo- and hypermethylated promoters, showing again that promoter-wide methylation does not correlate with expression.

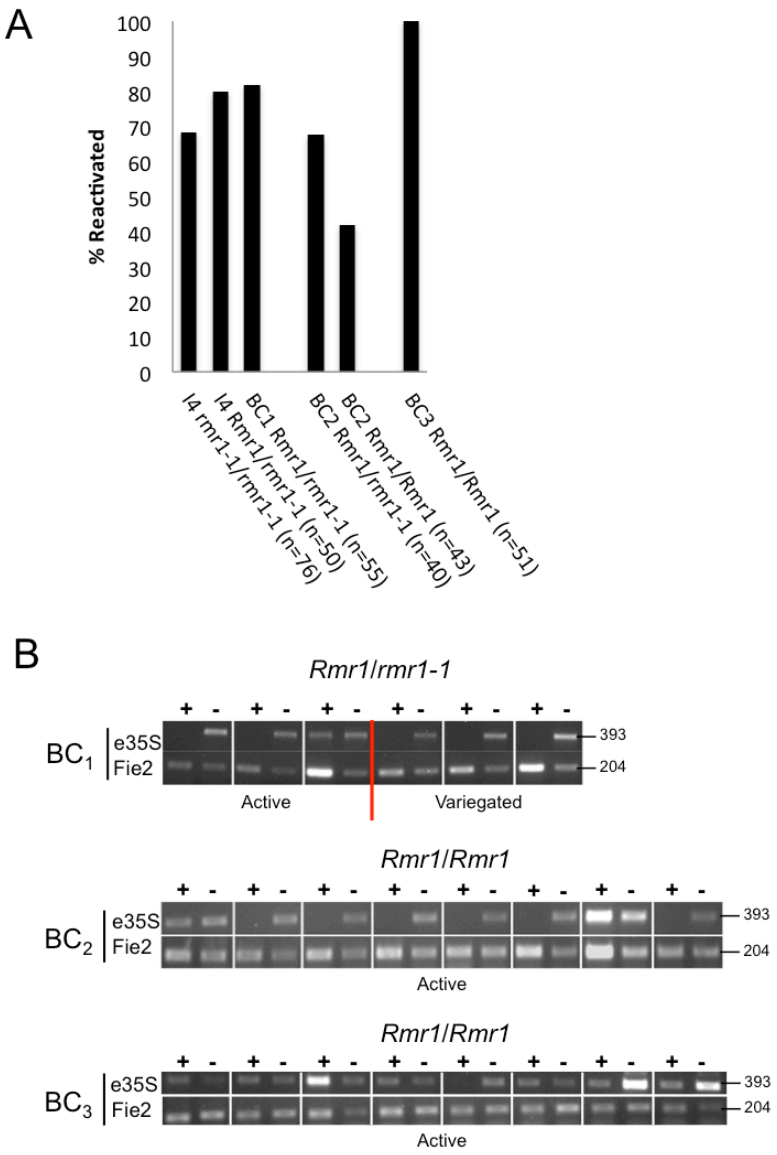


Figure 43: Percentage of reactivated NYR progeny and McrPCR methylation analysis of NYR-a* plants derived from *rmr1-1/rmr1-1* introgression. (A) Percentage of progeny reactivated following confocal microscopic analysis of the roots of newly germinated seeds in backcrossing of NYR-a* *rmr1-1/rmr1-1* with WT over three generations (BC_1, BC_2 and BC_3). (B) McrBC digestion and PCR amplification of the NYR e35S promoter and Fie2 (unmethylated control) in three NYR-a* and three NYR-v BC_1 *Rmr1/rmr1-1* individuals, eight NYR-a* BC_2 non-mutant and eight NYR-a* BC_3 non-mutant individuals. PCR product sizes (in bp) is indicated on right hand side of the panel.

4.3 Discussion

4.3.1 Detection of siRNAs targeting the NYR-v promoter

Small RNA libraries generated from NYR-v and non-transgenic samples display the now typical profile of an abundance of 21-25 nt small RNAs, with defined peaks at 22 and 24 nt. This profile represents the miRNA and siRNA small RNA classes present in maize, observed upon comparison of size and abundance (Nobuta et al. 2008). This profile is also observed in *Arabidopsis*, barley, rice and wheat with variation in the numbers of 21 and 22 nt small RNAs, but with the greatest abundance consistently belonging to the 24 nt size class (Henderson et al. 2006; Nobuta et al. 2008). Furthermore, the defined peak of 22 nt small RNAs is thought to be associated solely with maize (Nobuta et al. 2008). The predominantly 24 nt siRNAs detected in NYR-v were homologous to the e35S promoter and implicate the RdDM pathway heavily in the origin of silencing. However, without small RNA analysis of NYR-a derived from a separate transformation event it can only be assumed that the small RNAs identified are exclusive to NYR-v. The presence of 21 and 24 nt siRNAs are known to be involved the RdDM pathway, which mediates TGS (Mette et al. 2000; Simon and Meyers 2011). Small RNAs generated from tandem repeat regions are often of low abundance and therefore difficult to detect (Arteaga-Vazquez et al. 2010), and this agrees with the low siRNA abundance found at the NYR-v transgene.

Other studies have shown repeat regions or homologous sequence resulting in small RNA generation. Sequencing of small RNAs in tobacco lines with silenced transgenes showed alignments present only within regions of homologous sequence identity (Velten et al. 2012). Similarly, the two tandem elements of the NYR-v promoter share high sequence homology. Therefore the origin of siRNA in silencing NYR-v is likely due to the repetitive nature/homologous sequence of the promoter region and/or the proximity to the upstream TE. Numerous previous studies have shown similar attraction of methylation through RdDM and the presence of small RNAs (Matzke et al. 1989; Meyer et al. 1993; Velten et al. 2012; Khaitová et al. 2011). This also agrees with localisation of small RNAs within the NYR-v promoter, at the centre of the tandem elements, away from the distal unique region and also correlating with the presence of high levels of DNA methylation.

The RdDM pathway works primarily against TEs and repetitive elements, in addition to other DNA of foreign origin, such as transgenes, identified due to homologous/repetitive regions (Zhang et al. 2006). However, there have been studies of silenced transgenes that do not include repeat sequences, or homologous elements elsewhere in the genome (Pröls and Meyer 1992). These events are rare, and transgene silencing is most often driven by multiple copies of homologous sequences present in the genome, which significantly increases the chances of silencing (Meyer and Saedler

1996; Matzke et al. 2000).

4.3.2 RdDM mutants reactivate NYR-v

The RdDM mutants *mop1-1, rmr1-1, Mop2-1, rmr6-1* and *rmr2-1* were found to reactivate NYR-v. All of which have demonstrated an ability to reactivate silenced transgenes in maize, except *rmr6-1*, which as far as I am aware not tested (McGinnis et al. 2006; Sidorenko et al. 2009). These recent studies in maize have often taken advantage of fully silenced transgenes, such as BTG-s, observing that promoter methylation and its subsequent removal in mutant backgrounds is a requirement for silencing. The data I have gathered shows that NYR-v behaves differently to other published transgene systems and endogenous genes when introgressed into RdDM mutant backgrounds (McGinnis et al. 2006; Sidorenko et al. 2009; Sekhon and Chopra 2009). These include the observations that only a proportion of NYR-v plants are reactivated, promoter-wide methylation does not correlate with expression and an increase in the proportion of plants reactivated increases over multiple generations with some mutants, all of which will be discussed.

The known RdDM mutants utilised in this analysis, *mop1-1* (RDR2), *Mop2-1* (NRPD2/E2-like) and *rmr6-1* (NRPD1-like) represent the perturbing of Pol IV transcriptional and RNA-dependent RNA polymerase necessary for small RNA precursors, prior to their cleavage into 24 nt siRNAs by DCL3 (Law and Jacobsen 2010). This mechanism is thought to be the sole method of placing *de novo* methylation and the direction of asymmetric CHH methylation through targeting by small RNAs. These mutants present genome-wide crashes in the abundance of 24 nt siRNAs, although not a complete loss (Nobuta et al. 2008), and facilitate the loss of methylation from repetitive elements (Lisch et al. 2002). Furthermore, *rmr1-1* and *rmr2-1* also resulted in NYR-v reactivation. Only a little has been established of the function of each of these components, with the greater amount of focus on *rmr1-1*. Both are known to down regulate the abundance of 24 nt siRNAs in maize and effect the maintenance of DNA methylation at distinct loci (Barbour et al. 2012; Hale et al. 2007). Moreover, *rmr1* is thought to encode an SNF2-like protein, possibly granting a role in chromatin remodeling, but this is unconfirmed (Hale et al. 2007). Functional relationships between *rmr1*/*rmr2* and the known RdDM components (*mop1*, *mop2* and *rmr6*) also differs, despite similar involvement in siRNA accumulation. Such a difference exists as both *rmr1* and *rmr2* are not always required for paramutation in maize. Both are required for the maintenance of a paramutagenic state, but the absence of these components does not prevent paramutation of the *r1* or *pl1* loci (Barbour et al. 2012; Hale et al. 2007), whereas *mop1* and *mop2* are essential (Dorweiler et al. 2000; Sidorenko et al. 2009). Interestingly, *rmr2-1* causes a partial

loss of paramutation events at the *pl1* locus, but has no effect on paramutation of *r1* (Barbour et al. 2012). This difference in genome regulation shown by the mutants may indicate overlapping functions as well as distinct functions for these two sets of components. Furthermore, investigation into where *rmr1* functions in RdDM has lead to the proposal that it interacts upstream of *mop1* and may also interact with other RDR proteins, such as RDR6 (Hale et al. 2009).

Comparison with work conducted in *Arabidopsis* shows similarity in function of RdDM components, with mutants of RDR2 and Pol IV resulting in loss of 24 nt siRNAs and corresponding maintained DNA methylation (Nobuta et al. 2008). However, much of the studies conducted have not observed individual silencing events, and therefore may neglect the consequences of locus/sequence specific silencing, as observed in studies in maize. One study in *Arabidopsis* utilises a methylated and silenced *FWA* transgene demonstrating the requirement of the known RdDM components in gene silencing (Greenberg et al. 2011; Pikaard et al. 2008). This investigation also utilises further downstream RdDM mutants currently not available in maize, such as CLSY1 and DRD1, finding that they are as essential as the upstream Pol IV elements. My investigation also utilised *rmr1-1* and *rmr2-1* mutants in maize which currently have no known orthologues in *Arabidopsis*, however, SHH1 is thought to be a good candidate, associating with the Pol IV complex (Hollick 2012). It has been shown to be required for the maintenance of *de novo* methylation and the accumulation of siRNAs (Law et al. 2011). Other investigations have focused on expression and methylation state changes of endogenous repetitive elements regulated by RdDM, but greater analysis has been conducted in maize with regards to the effects of expression in mutant backgrounds and also expression following mutant segregation (McGinnis et al. 2006; Madzima et al. 2011; Sekhon and Chopra 2009; Woodhouse et al. 2006). Therefore, the approach taken more commonly in maize may reveal more about locus/sequence specific silencing as often genes can behave differently in mutant backgrounds.

Other mutants including, *Ufo1-1*, *mom1-like* and *Rpd3-like* were also utilised in my investigation of NYR-v silencing, failing to result in reactivation. This also serves as an investigation into the role of *ufo1* and if it behaves within the RdDM pathway. Despite *ufo1* possessing a role in maintaining silencing in *b1* paramutation (Sekhon et al. 2012), which is shared with *mop1, rmr1, mop2* and *rmr6*, no reactivation of the transgene was detected. It has also been shown to down regulate the implementation of repressive histone modification H3K9me2 at the *P1-wr* allele suggesting a role in multiple epigenetic pathways (Sekhon et al. 2012; Chopra et al. 2003). These studies have only investigated how paramutation is affected by this mutant and not traditional gene silencing of transgenes and endogenous elements. Therefore, it is difficult to discern if gene regulation can

depend upon its action in maize, but NYR-v appears to be unaffected by its loss of function. Similarly, *mom1-like* and *Rpd3-like* also failed to reactivate the NYR-v transgene, despite previous studies in *Arabidopsis* and maize showing release of silencing of methylated genes (Amedeo et al. 2000; Rossi et al. 2003). *Rpd3-like* encodes a histone deacetylase shown to repress transcriptional activity of a GUS reporter, however, the focus of the report was in determining if there was cooperation between the products of *Rpd3-like* and another gene involved in silencing (Rossi et al. 2003). Therefore, the extent to which *Rpd3-like* contributes in gene silencing is unknown. MOM1 has been shown to regulate TGS in a methylation independent way and is currently thought to associate with Pol V in RdDM, attracting repressive histone modifications (Zhou et al. 2010; Numa et al. 2009). No analysis with *mom1-like* or *Rpd3-like* has been conducted in maize, therefore it is difficult to draw a firm conclusion from the lack of activity when combined with NYR-v. These mutants appear to not be associated with the RdDM components that have presented reactivation of NYR-v.

Interestingly heterozygous RdDM mutants (*mop1-1, rmr1-1, Mop2-1,* and *rmr6-1*) also showed reactivated progenies from the I_2 generation, which may indicate a gametophytic effect in the expression of NYR in the following generation. Reactivation may therefore be dependent on the RdDM environment of the forming male and/or female gametes for expression in the next generation. During gametogenesis in plants, meiosis occurs generating haploid micro- and megaspores, which then require further postmeiotic divisions to form the sperm and egg (Takeda and Paszkowski 2006). A study in *Arabidopsis* with *Met1/met1-1* heterozygous mutants resulted in hypomethylated plants, leading to the hypothesis of demethylation occuring within the gametes as a result of the mutation in a haploid environment (Kankel et al. 2003). This has subsequently been reinforced by *Met1/met1-1* showing hypomethylation and release of TGS in the following generation (Saze et al. 2003). Therefore, the RdDM mutants tested in the analysis of NYR-v may have performed by a similar action, resulting in reactivation within the heterozygotes. This also potentially implicates the epigenetic state of the gamete(s) in the regulaton of gene expression in the mature plant.

4.3.3 RdDM mutants reactivate NYR-v at different potencies

Of the RdDM mutants leading to the release of NYR-v silencing, varying percentages of progeny were reactivated and others remained silenced/variegated. Previous work with *rmr1-1, mop1-1* and *rmr2-1* has shown efficient restoration of a transgene expression after a single mutant introgression (McGinnis et al. 2006). In this study, silenced 35SBTG, present in 2 copies, were reactivated upon introgression with the mutants showing a reduction in methylation from the promoter. Therefore,

the conclusion can be made with certainty that methylation and expression correlated, controlled by RdDM, a conclusion which cannot be drawn for NYR-v. My results, however, indicate that reactivation of NYR-v is possible but inefficient in these mutants, with a proportion of offspring showing reactivation. This phenomena has been observed in the reactivation of silenced BTG from the putative *tgr2* (*transgene reactivated 2*) mutant, revealing a lower than expected proportion of active progeny and also producing plants with sectored/variegated transgene reporter expression (Madzima et al. 2011). The role of *tgr2* is unknown and studies utilising known RdDM components more often than not result in complete reactivation.

Furthermore, the percentage of reactivated progeny differing between the RdDM mutants tested hints at the importance of different components for silencing. For instance *rmr1-1* and *rmr2-1* exhibited the highest percentages of progeny reactivated, and as previously stated, their functions remain unknown, but their loss results in the reduction of siRNAs (Hale et al. 2007). The SNF2-like protein encoded by *Rmr1* is also thought to interact with chromatin marks, but little else is known about its function. Previous work has shown that release of silencing of the *35S:NPTII* transgene in *Arabidopsis* has been achieved to greater effect with SMZ, which reduces DNA methylation and H3K9me2, over DNA-methylation inhibition alone (Zhang et al. 2012). Therefore, repressive histone modifications may have a role in NYR-v silencing. This example displays the relationship that can exist between DNA methylation and repressive histone modifications in the genome, providing TGS.

The mutants with direct roles in 24 nt siRNA biogenesis *mop1-1, Mop2-1* and *rmr6-1*, however, reveal lower proportions of reactivated progeny. This is surprising as the presence of 24 nt siRNAs and DNA methylation at the promoter suggested that the RdDM pathway was the cause of silencing. Interestingly, CHH methylation has remained at the promoter of NYR-a* plants despite their RdDM mutant backgrounds. The current model of RdDM proposes that Pol IV and RDR2 components are essential for the generation of 24 nt siRNAs and therefore the targeting of DRM2 to the genome (Law and Jacobsen 2010). However, recent work conducted in *Arabidopsis* has shown the maintenance of CHH methylation in Pol IV and V double mutants (Zhong et al. 2012; Wierzbicki et al. 2012). The Wierzbicki et al. (2012) study also shows that although CHH methylation is lost at 50-60% of cytosine locations in the genome, new CHH methylation, not present in the WT background, is placed within pericentromeric regions. Therefore, the targeting of methylation to the NYR promoter may be the result of other unknown factors than Pol IV derived siRNAs alone as promoter-wide methylation remains (at lower levels than NYR-v in a WT background). Alternatively, histone modifications can also attract the action of methyltrans-

ferases, allowing for the maintenance of methylation without siRNA signalling (Law and Jacobsen 2010). The loss of function of histone methyltransfersase KYP, responsible for H3K9me2, results in a substantial reduction in DNA methylation, with others such as SUVH5 and 6 identified (Ebbs and Bender 2006; Ebbs et al. 2005; Jackson et al. 2002). My investigation has found that NYR-v silencing is more heavily dependent on the functions of *rmr1* and *rmr2* than RdDM components directly involved in the production of siRNAs due to the different proportions of reactivated plants achieved. However, the reasons for this remain unclear due to the lack of information concerning the function of both of these components.

4.3.4 Loss of NYR-v silencing in RdDM mutants is progressive over multiple generations

Analysis of *mop1-1* and *rmr1-1* mutants over three generations revealed that proportions of reactivated progeny increased progressively. Similarly, previous work by Woodhouse et al. (2006) in maize has shown that progressive reactivation of the transposable element *mudrA* occurs within a *mop1-1* mutant background over subsequent generations. The effect has also been observed in *Arabidopsis* utilising a *ddm1* mutant background (Jeddeloh et al. 1999). This fits with the theory that methylation is lost gradually over multiple generations, releasing an area from heterochromatin and silencing (Sekhon and Chopra 2009). It is believed that the introduction of mutants involved can present a 'slow but cumulative decay of epigenetic states' (Sekhon and Chopra 2009; Woodhouse et al. 2006). The loss of a single component within a pathway may set this decline in motion due to multiple components/pathways involved in silencing. The increase of NYR-v reactivation in *mop1-1* and *rmr1-1* backgrounds may be a result of this, whereby the gradual loss of DNA methylation due to depletion of siRNAs. However, promoter-wide methylation at NYR does not correlate with expression, as shown by hypomethylated profiles of both NYR-a* and NYR-v homozygous mutants between I_2 and I_4 generations. This suggests that the deteriorating epigenetic mark may be a specific site of methylation or perhaps the change in repressive histone modifications.

Interestingly, the Sekhon and Chopra (2009) study, which investigated *Ufo1* inducing hypomethylation of the *P1-wr* allele in the maize pericarp, showed gradual loss of methylation over multiple generations in the form of an increasing number of unmethylated clones compared to hypermethylated clones sequenced. This gradual change has also been observed in endogenous genes, such as epimutations forming within *fwa* due to multiple generations in a *ddm1* mutant background in *Arabidopsis* (Soppe et al. 2000). These results demonstrate the correlation between DNA methy-

lation and expression in the decreasing number of clones with high methylation and the increasing proportion of cells exhibiting expression. My data shows a more uniform reduction in methylation at all the NYR-a* clones sequenced. Therefore, variation in proportions of plants observed appears to have more common with observations by Woodhouse et al. (2006) , rather than Sekhon and Chopra (2009) as it is a proportion of the progeny exhibiting different epialleles and not a changing epigenetic landscape within each individual. However, it is unclear what epigenetic mark is deteriorating across generations for the increase in NYR-a* plants as a correlation between promoter-wide methylation and expression was not drawn. Therefore, as previously, this suggests that methylation at a different location or he specific loss of methylation at vital cytosines (to be discussed) are responsible for silencing, providing that RdDM is the sole component.

Contrastingly, *Mop2-1* did not behave in the same manner as *mop1-1* and *rmr1-1*, instead maintaining a lower proportion of NYR-a* progeny uniformly over three generations (I_2 to I_4). Mutation of *mop2* results in a genome-wide reduction in siRNAs and regulates gene silencing as a subunit of Pol IV in the the RdDM pathway (Sidorenko et al. 2009). Published work has shown *Mop2-1* to behave in the same manner as other RdDM mutants, such as *mop1-1*, in the reactivation of silenced transgenes, through the removal of methylation (Sidorenko et al. 2009; Stonaker et al. 2009). Therefore, the difference in reactivation proportions compared with *mop1-1* are confusing and the only likely conclusion is that different Pol IV isoforms may act redundantly, or act on different silencing events, dependent on unknown factors (Stonaker et al. 2009). This is also supported by the observation that NRPD2/E2-like (*Mop2-1*) deficient plants do not exhibit as severe adverse phenotypic results as NRPD1 (*mop1-1*) deficient plants, possibly due to redundancy between the three NRPD2/E2-like genes identified in maize (Pikaard and Tucker 2009). Extra functionality may therefore be able to maintain a level of silencing at NYR-v in the absence of *mop2*.

4.3.5 Loss of methylation in the *asf-1* region of the NYR-a* promoter

Bisulfite sequencing analysis of the promoter region of NYR-a* in RdDM homozygous mutant backgrounds revealed a uniform drop in methylation across all sequence contexts, but not a complete loss, compared with NYR-v in a WT background. This is in contrast to other research which has showed complete loss of methylation upon transgene reactivation at the promoter region (Madzima et al. 2011). Maintenance of methylation must therefore be somewhat independent of the RdDM pathway as it has remained meiotically heritable over multiple generations, as has previously been discussed. However, the reason for differential expression still remains, not explained by the uniform drop in methylation at the promoter, but perhaps due to the loss of methylation

at strategically important cytosine positions.

This chapter has shown that there is no correlation between promoter-wide methylation and expression of NYR that has been detected. Bisulfite sequencing of NYR-a* and NYR-v plants both homozygous for an RdDM mutant would have therefore been best to obtain this information. Differences in methylation detected (if any) would indicate sites of importance in determining the role of DNA methylation in silencing. The importance of the methylation at the *asf-1* region was discussed in the previous chapter, stating that two CG locations within the *asf-1* region are responsible for controlling transcriptional activity (Kanazawa et al. 2007a). Both of these sites are found with greatly reduced methylation levels in NYR-a* with RdDM mutant backgrounds than NYR-v in a WT background. Prior work has shown that loss of methylation at *asf-1* returns functionality to the promoter (Meyer et al. 1994).

Furthermore, previous work carried out with MOM, demonstrated transcription of a heavily methylated region, formerly silenced (Amedeo et al. 2000). Perturbing MOM function appeared to lead to a release from silencing, while high levels of methylation remained. The conclusion generated from this study was that transcription and DNA methylation could be independent of one another. However, small amounts of methylation at distinct positions can also have a great amount of control in determining transcription, which would not have been detected in this study (Kanazawa et al. 2007a). Again, this adds weight to the argument that promoter wide methylation is not responsible for silencing, but rather individual positions. This is reinforced by the small amount of methylation possessed by variegated/sectored BTG plants compared to the completely hypomethylated silenced BTG plants in putative *tgr2* mutant backgrounds (Madzima et al. 2011). It remains unclear if methylation alone at the *asf-1* region is the sole reason for NYR-v silencing due to the remnants of methylation present at NYR-a* in RdDM mutant backgrounds and the lack of comparison with NYR-v plants in the same background.

4.3.6 Detection of L-UPN readthrough transcript derived upstream of NYR-a*

Interestingly, alongside an increase in YFP reporter expression following mutant introgression of NYR-v, mRNA was also detected in the form of the L-UPN transcript. cDNA synthesis of which relied on the use of oligo dT18, indicating that the RNA amplified was polyadenylated. There is no transgene promoter upstream of this location suggesting that transcription is initiated from within the *gypsy*-like TE flanking sequence, which may become transcriptionally active due to loss of RdDM. DNA methylation is known to be responsible for the silencing of unfavourable genic regions, such as repetitive regions and TEs and the release of which has shown to cause changes

in the expression (both up and down) of TEs (Jia et al. 2009). Therefore, loss of RdDM could result in relaxation of chromatin and transcription from within the repetitive region, subsequently spreading into the transgene (Eichten et al. 2012; Singh et al. 2008). L-UPN transcript RNA detected also correlates with increased YFP reporter expression from NYR-a* reactivated plants and not present within the NYR-v, suggesting that its activity changes with NYR expression. If this is the case, it would indicate that a position effect may be responsible for silencing of NYR, whose expression is dependent upon the epigenetic landscape and transcriptional state of the surrounding endogenous DNA. However, as previously this remains unclear without comparison of NYR-a* and NYR-v both derived from an RdDM homozygous mutant, as the detected RNA may result from the homozygous mutant background and not the transgene expression. These data points to the possible influence of neighbouring sequences, in a position-effect manner, on NYR-v expression.

4.3.7 Restoration of *rmr1* does not reestablish NYR silencing

Introgression of NYR-v into a *rmr1-1* mutant background and then restoring *rmr1* showed that an NYR-a* state is heritable across multiple generations. My results are in contrast with McGinnis et al. (2006), which shows silencing is efficiently reinstated from transcriptionally active BTG *rmr1-1* derived lines. Interestingly, this study also observes that BTG reactivation remained following outcrossing from *mop1-1* and *rmr2-1*. Therefore, the failure to reestablish silencing at NYR following restoration of *rmr1* may the be the result of a changed chromatin environment resisting silencing. Furthermore, this trait has been observed in *Arabidopsis,* where introgression of the *FWA* transgene into a *ddm1* mutant background is also associated with retaining mutant expression levels despite segregation out of the *ddm1* background (Stokes and Richards 2002). However, a study also in *Arabidopsis* has found delayed resilencing following outcrossing from TGS mutants, proposing that a gradual resilencing process was required (Scheid et al. 1998). This suggests that the proportion of plants showing NYR reactivation would decrease over multiple generations, but this not observed. Strikingly the BC_3 generation, the second generation of NYR-a* in a non-mutant background, all progeny observed were reactivated. It appears likely that heritable transmission of the NYR-a* epiallele is due to selection for this epiallele through the BC generations. This is similar to the description of 'genetic assimilation', whereby the development of a specific trait can be induced and selectively propagated until the trait can be displayed without the presence of the inducer (McGinnis et al. 2006). Work conducted by Waddington (1953) proposes this upon observation of environmental stress inducible traits studies in *Drosophila*. The adoption of NYR-a* may be a result of a similar process through alteration of DNA methylation and/or

chromatin introducing a stable epiallele into all progeny. McGinnis et al. (2006) conclude that *mop1-1* and *rmr2-1* mutants allow for a change in chromatin structure that can be maintained meiotically following backcrossing to WT, and that *rmr1-1* does not. The opposing data that I have generated may therefore indicate that the effects of silencing are locus and sequence specific.

Interestingly, DNA methylation analysis of outcrossed NYR-a* non-mutant plants revealed that despite restoration of RdDM, methylation was not always returned to WT levels in the promoter region. BC_2 exhibited that 25% of plants retained hypomethylation (observed within the mutant introgression) and in BC_3this figure was 88% of plants. As previously this suggests that promoter-wide methylation is not responsible for NYR-v expression. Currently, of the studies that have investigated this effect, there has not been thorough investigation of the progenies to determine the changes implemented to an allele altered in expression by an RdDM mutant background.

4.3.8 Summary

In summary, RdDM mutants were found to reactivate NYR-v, producing different proportions of reactivated plants. Multiple generations of introgression into RdDM mutant backgrounds also showed loss of silencing to be progressive suggesting deteriorating epigenetic marks.

Loss of silencing was found not to correlate with promoter-wide DNA methylation but may be due to the reduction of methylation at the 3' region of NYR-a* plants.

Subsequent, restoration of RdDM function found that following mutation of *rmr1*, restoration was not sufficient to reestablish NYR silencing.

5 Assessment of *trans*-silencing in Vegetative and Reproductive Tissues Utilising NYR-v

5.1 Introduction

5.1.1 Chapter aims

The previous chapter has shown that the NYR-v is maintained meiotically through multiple generations in a WT background, therefore it may be used as a tool in studying paramutation-like effects. Currently, true paramutation has only been observed in endogenous genes and transgenes that incorporate endogenous sequences already known to be paramutagenic. Other transgenes have only shown *trans*-silencing (see 1.5.2) events and not subsequent paramutagenic activity following removal of the original paramutagenic signal origin. The determinants with regards to structure, sequence and location of genes, which are paramutagenic, have not yet been finalised. Therefore, investigation with NYR-v may be able to demonstrate true paramutation-like events in transgenes.

Moreover, NYR-v can potentially be used as a tool in the investigation of the movement of small RNAs between seed components. DNA methylation patterns presently suggest that seed components can communicate, transferring epigenetic states, possibly mediated by the action of small RNAs. The presence of 24 nt siRNAs and the possibility of a *trans* silencing ability of transgenes with sequence homology means that this communication can be tested providing there is a silencing element (NYR-v) and a target element (transgene with sequence homology), but only one within each seed component. The ability to achieve this is explained by heterofertilisation events in maize (see 5.1.5).

This chapter included two objectives:

- To determine if NYR-v can act in *trans* and if it has paramutagenic activity.
- To assess if NYR-v can be used as a tool to demonstrate the movement of small RNAs between seed components.

In order to describe how epigenetic information may be passed between seed components it is important to understand how small RNAs have been shown/hypothesised to move between tissues. My experiments aim to demonstrate this through the pairing of genetically different endosperm and embryo tissues. It is therefore important to understand how genetically distinct endosperm and embryo seed components can be generated through heterofertilisation in maize and identified by anthocyanin reporter expression, which is discussed in the following sections.

5.1.2 Anthocyanin pigment expression in the study of epigenetics

The anthocyanin pathway in maize produces a deep red/purple pigmentation in most vegetative tissues and has been greatly exploited in the study of genetics and epigenetics, providing a clear visual marker of gene expression (Fig.44). It is considered non-essential in maize, but the pigment is abundant in the plant kingdom, serving as protection against UV radiation and the attraction of pollinators (Petroni et al. 2000). The study of this pathway has led to the discovery of paramutation (Brink 1956) and genomic imprinting (Kermicle 1970) in maize, utilsing the *red1* (*r1*) gene, a transcription factor involved in anthocyanin biosynthesis.

In addition, the *booster1* (*b1*) gene, another transcription factor that regulates anthocyanin in maize has been widely used, also exhibiting paramutation (Coe Jr 1959). Moreover, it is the preferred gene for incorporation into transgenes as a colour reporter expression in maize, activating the anthocyanin biosynthesis pathway in most tissues (McGinnis et al. 2006; Madzima et al. 2011; Goff et al. 1990).

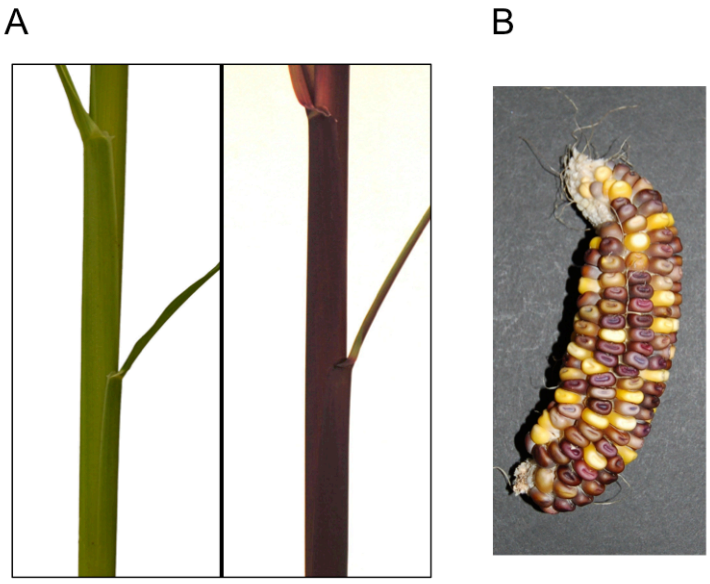


Figure 44: Anthocyanin colour pigment in maize. Images of stems (A) and seeds attached to ear (B) with and without anthocyanin pigment (purple colour) expression.

5.1.3 Intercellular Small RNA Movement

It has long been theorised that small RNA molecules can communicate epigenetic states between different tissues in plants, in a similar manner to how hormones operate in cell differentiation (Carlsbecker et al. 2010). The mobility of silencing signals has been previously shown through the use of an Inverted Repeat (IR) transgene reporter inducing silencing of a target reporter across

grafting junctions or by specific expression within companion cells in *Arabidopsis* (reviewed by Brosnan and Voinnet 2011). Two studies have also shown that endogenous small RNAs can be mobile, utilising root/shoot grafting experiments in *Arabidopsis* (Molnar et al. 2010; Dunoyer et al. 2010a). The first study found that small RNAs could move via the phloem, in a source-sink direction initiating silencing. In addition, the detection of total small RNAs by deep-sequencing in grafted root tissue was compared between WT and dicer-*like* 2,3 and 4 (*dcl2,3,4*) triple mutant shoots (Molnar et al. 2010). The mutant was unable to produce 22-24 nt small RNAs from larger RNA precursors. The results of which showed a build up of small RNAs into the *dcl2,3,4* mutant graft indicating movement from the WT roots (Molnar et al. 2010). The second study observed the activity of two endogenous IRs, showing that small RNAs (including 24 nt size class) were mobile and could also alter methylation at distant locations, specifically over a graft junction (Dunoyer et al. 2010a). A further study also showed this by silencing a GFP reporter transgene in shoots with an IR transgene present in root tissue (Melnyk et al. 2011) (Fig.45A). These studies demonstrate how endogenous genes may be regulated by mobile small RNAs in vegetative tissues, but have lead to the hypothesis that small RNAs and the inter-cellular communication of epigenetic marks in reproductive tissues and could have a role transgenerationally.

This hypothesis has been reinforced by the detection of movement of small RNAs between the vegetative pollen nucleus and sperm cells (Slotkin et al. 2009) (Fig.45B). DCL3 derived small RNAs generated as a result of active demethylation and TE reactivation within the vegetative nucleus were subsequently detected in sperm cells. DME-deficient companion cells were also shown to produce gametes with lower levels of CHH methylation at TEs (Ibarra et al. 2012). In addition, a similar situation is observed between the polar nuclei of the central cell and the egg cell (Ibarra et al. 2012) (Fig.45C). Active demethylation by DME of the central cell in *Arabidopsis* is thought to produce the same result as the demethylation of the vegetative nucleus in pollen. The same report showed that a microRNA expressed exclusively in the central cell could silence a transgene GFP reporter present in the egg. This suggests that the companion cells of reproductive tissues reinforce TE methylation and silencing in plant gametes through contribution of small RNAs. However, this is not unique to plants and has also been hypothesised in other animals, such as *Danio rerio* (zebrafish) (Jiang et al. 2013).

Finally, small RNA movement has also been hypothesised between the endosperm and embryo tissues of a developing seed (Mosher and Melnyk 2010) (Fig.45D). These are the two products of double fertilisation in formation of a mature seed (see 1.2.3). Methylation analysis of seed tissue in *Arabidopsis* revealed genome-wide demethylation in the endosperm compared with the embryo (or

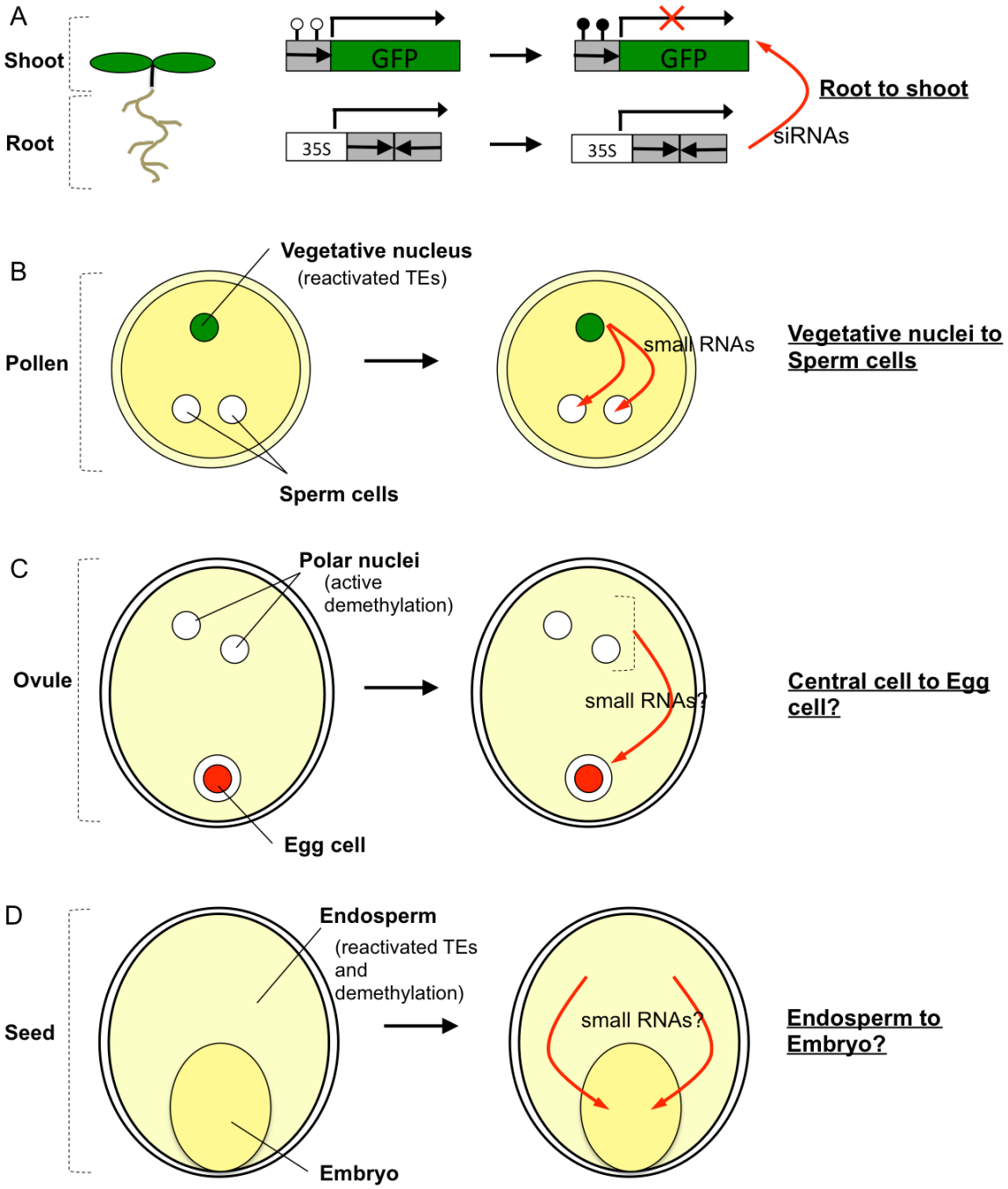


Figure 45: Diagram illustrating the movement/proposed movement of small RNAs between different tissues. (A) Movement of small RNAs between genetically dissimilar root and shoots that have been grafted together, inducing TGS and DNA methylation of target GFP reporter transgene (Melnyk et al. 2011). (B) Accumulation of small RNA in sperm cells derived from reactivated TEs in the vegetative nucleus (Slotkin et al. 2009). (C) Proposed movement of small RNAs between the central cell and the egg cell in ovules due to demethylation of central cells and reinforcement of methylation in gametes (Ibarra et al. 2012). (D) Proposed movement of small RNAs between endosperm and embryo tissue of developing seeds (Mosher and Melnyk 2010).

vegetative tissue) (Hsieh et al. 2009). Although not as severe, a loss in methylation is also observed in maize endosperm (Lauria et al. 2004). The result is a loss of regulation of the genome, TEs are activated and previously epigenetically silenced regions are released. Transcriptome analysis of developing maize seeds has found that DCL, AGO and RDR genes required for RdDM are upregulated in the embryo compared with the endosperm, which appears to form the opposite picture of operations to *Arabidopsis* (Lu et al. 2013).

The transfer of epigenetic states inferred and shown by elements of these reproductive tissues suggests that epigenetic information can be transmitted to the next generation, perhaps serving as an environmental sensor to aid fitness of offspring (Mosher and Melnyk 2010). Work conducted in *Caenorhabditis elegans* has already shown that piwi-interacting RNAs (piRNAs) can be transgenerationally inherited through the germline without requiring the initial small RNA stimulus (Ashe et al. 2012) and that the Argonaute, HERITABLE RNAi DEFECTIVE 1 (HRDE 1), allows for small RNA modification of germ cell nuclei (Buckley et al. 2012). Together, these studies show how small RNA contributions to the germline can produce transgenerational memory, and may transmit information to future generations, which has similarities to the Lamarckian evolutionary theory.

5.1.4 Mechanisms of small RNA movement

The mechanisms allowing small RNA molecules to be exported, traverse cells and imported are largely unknown. Research conducted into movement in plants has shown that RNA can move between cells via the plasmodesmata and the vascular tissue (Melnyk et al. 2011) and in mammals miRNAs have been detected in the blood stream (Hunter et al. 2008). Therefore this presents two potential modes of movement in an organism; local and systemic.

Locally, RNA molecules may be able to move independently between cells, as observed in SULPHUR (SUL) silencing by transgene IR elements with a phloem specific promoter (Dunoyer et al. 2005). SUL is required for photo-protection and is ubiquitously expressed. Silencing was observed only a limited distance from the vascular tissue in the form of photobleaching of the leaf tissue. Phloem-based movement for larger distances has also been shown, utilising *in vivo* phloem tracers (Tournier et al. 2006). This agrees with findings showing that the source to sink direction of silencing is most effective (Melnyk et al. 2011). Larger distances may also require specialised transportation, microvesicles have been found associated with RNA molecules in mammalian blood (Hunter et al. 2008), shown to ferry miRNAs systemically in humans resulting in gene silencing (Xiao et al. 2007). Also, the initial secretion of RNA molecules is thought to occur passively

and actively, passively in the breakdown of a cell and actively through secretion by microvesicle or an RNA-binding protein, such as a high-density lipoprotein (HDL) (Chen et al. 2012). The subsequent uptake of RNA molecules however, is an area which requires more study. SID-1 is a known RNA transporter in *C. elegans* that is required for the active import of RNA, however, there is much about the pathway that is currently unknown (Jose et al. 2011). It is thought that the delivery of RNA molecules to recipient cells works by endocytosis, phagocytosis or fusion with the plasma membrane (Chen et al. 2012), but again it is unknown whether specific transmission vehicles recognise specific cell types for example. Artificial systems have also unearthed the ability for small peptide chains to associate with RNA molecules and deliver them into cells, as explored successfully in RNAi experiments using a synthesised 12mer arginine peptide (Unnamalai et al. 2004). In addition, the question of what is transmitted in inter-cellular RNA silencing has also been investigated. Whether it is a dsRNA precusor, an intermediate or a diced mature RNA molecule, such as a 24 nt siRNA. A study in *Arabidopsis* has revealed that it is the processed RNA molecule that is mobile, through the use of P19 tombusviral suppressor of RNA silencing (VSR) which binds 21 nt small RNAs, which arrested distant RNA induced silencing (Dunoyer et al. 2010b).

5.1.5 Heterofertilisation

In order to produce genetically distinct embryo and endosperm tissues within the seed for analysis of small RNA movement, a low frequency event called heterofertilisation can be employed forming such seeds. This can be due to fertilisation of the egg cell and central cell by sperm cells of different pollen grains, or, conversely due to differing egg cells and central cell forming, fertilised by the same pollen (Gao et al. 2011). The phenomenon was first studied by Sprague (1929) in maize, which remains the only organism in which it has been documented. Subsequent studies carried out by Sprague (1932) and Sarkar and Coe (1971) provide average frequencies of 1.5% and 1.15% respectively of a heterofertilisation event occurring. However, frequencies as high as 5% has been shown by Robertson (1984) and 25% was reported by a single line in initial work by Sprague (1932). More recently, an experiment utilising heterofertilisation to look into resource allocation in the seed recorded a frequency of 0.58% (Wu et al. 2013). These studies have utilised colour pigment production in the aleurone and the scutellum as defining factors for the genotype of endosperm and embryo tissues. Plants heterozygous for colour genes involved in anthocyanin or carotenoid biosynthesis were crossed into non-coloured lines and a heterofertilisation event is observed as a mismatch between aleurone and scutellum pigment (Sprague 1932; Sarkar and Coe 1971; Robertson 1984). Fig.46 provides a schematic showing a cross section of the maize seed,

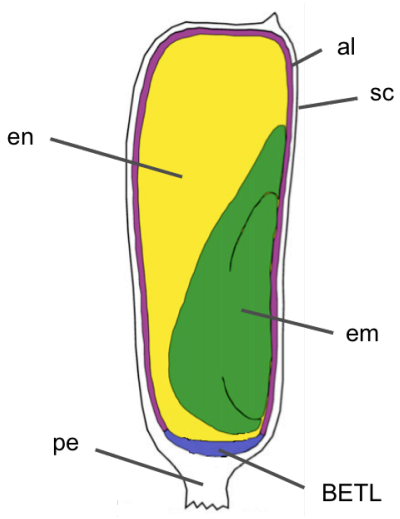


Figure 46: Schematic diagram cross-section of a maize seed. Green, embryo (em) (mostly the scutellum); yellow, starchy endosperm (en); purple, aleurone (al); blue, basal endosperm transfer layer (BETL); white, maternal tissue – pedicel (pe) and seed coat (sc). Adapted from Becraft and Gutierrez-Marcos (2012)

indicating the location of the aleurone and scutellum. Work conducted by Sprague (1932) observed the aleurone and scutellum pigment of 13,716 seeds derived from 59 independent crosses to find heterofertilisation events, a further 719 seeds from 7 independent crosses lead to the finding of a line showing 25% heterofertilised seed. Sarkar and Coe (1971) analysed 8,430 seeds, selecting only seeds with colourless endosperm/aluerone, seeking seeds with coloured scutellums. The Robertson (1984) study looked in greater detail at the appearance of the earlier 25% frequency (Sprague 1932) and also how frequency was affected by the use of differing parent lines. Some 64,615 seeds were analysed across these lines, showing an average of 0.8% (0.0% - 5.0%) heterofertilisation. The must recent study conducted using this system observed pigmentation in 40,638 seeds derived from 103 crosses showing an average frequency of 0.58% (Wu et al. 2013).

Recent studies into heterofertilisation frequency have also used molecular markers in its detection, to derive greater accuracy (Gao et al. 2011). This is because the use of colour pigments in seeds is known to be affected by environmental conditions, which may influence frequency results (Ford 2000). The Gao et al. (2011) study estimated an average heterofertilisation frequency of between 0.14% and 3.12% (average of 1.46%), consistent with estimates made in the previous studies, determining the genotype of endosperm and leaf of the plant generated. This study analysed an average of 692 seeds from each of 11 independent cross events. The frequency of events is thought to range between 0.5-5%, and currently dependent on unknown factors.

5.2 Results

This results chapter has been split into three parts. The first part addresses the study of NYR-v in *trans*-silencing of other transgenes. The second part explores the paramutation-like effects of NYR-v. In the final part I investigate the use of NYR-v in the study of small RNA movement in the maize seed.

5.2.1 Homology-dependent NYR-v *trans*-silencing

In order to assess if silencing of NYR-v can be transferred in *trans* via sequence homology, experiments were performed combining NYR-v and a second transgene, BTG (b1 genomic transgene - referred to here as BTG-active (BTG-a)). The BTG-a promoter shares 348 bp sequence homology (which includes 13 mismatches) to the NYR-v promoter, and drives expression of *b1* resulting in anthocyanin pigment production (Fig.47). Alignment of promoter sequences show that small RNA sequences previously discovered in NYR-v plants match in regions of homology with the BTG-a promoter (Fig.48B). Three identified small RNAs (24 nt, 23 nt and 24 nt in size) match with complete homologous sequence identity between the NYR-v and BTG-a promoters, whereas another three small RNAs (24 nt, 21 nt and 24 nt in size) have mismatched sequence identity with the BTG-a promoter. The latter small RNAs align to regions a greater distance from the TSS and other functional elements.

This analysis indicates that small RNAs derived from NYR-v could potentially induce *trans*-silencing of BTG-a.

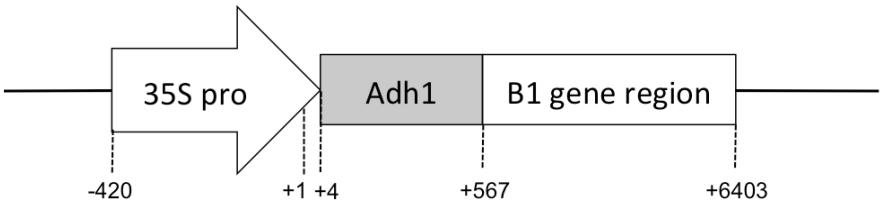


Figure 47: Schematic diagram of b1 genomic transgene (BTG-a). 35S CaMV promoter driving expressing of the B-I allele of the b1 gene, part of the anthocyanin pigment production pathway. CaMV, Cauliflower Mosaic Virus; Adh1, alcohol dehydrogenase1. Positions relative to transcriptional start site (+1). Adapted from McGinnis et al. (2006).

5.2.2 Phenotypic analysis of BTG-a after exposure to NYR-v

To this end, hemizygous NYR-v and BTG-a transgenic plants were crossed as shown in Fig.49, generating both separate hemizygous NYR-v and BTG-a progeny, combined progeny and progeny

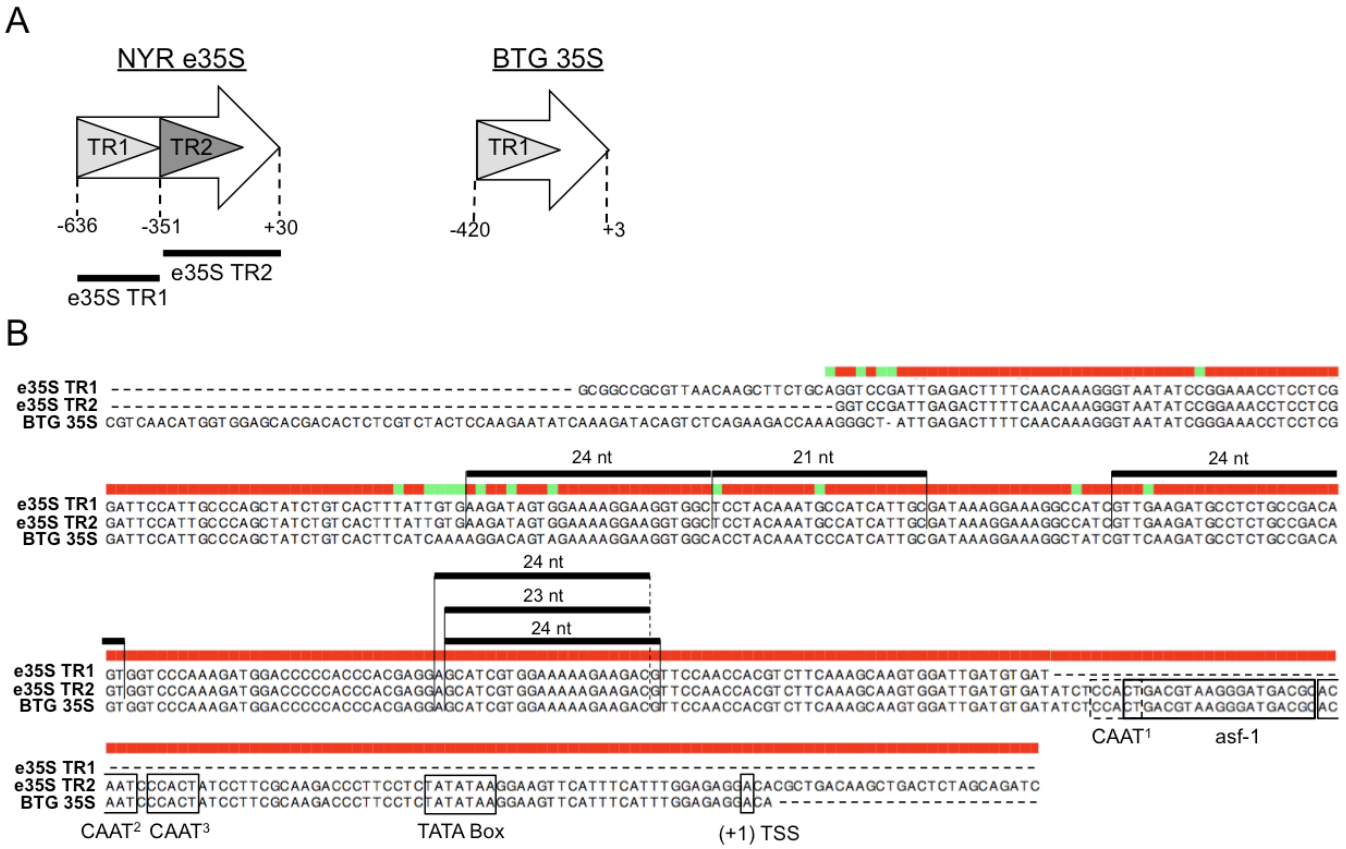


Figure 48: Sequence analysis of NYR-v e35S and BTG-a 35S promoters. (A) Schematic diagram of promoter regions and positions of repeats. (B) ClustalW alignment of NYR-v e35s (separated) and BTG-a 35S promoter sequences, red showing homology and green showing a mismatch between the NYR-v tandems and BTG-a promoters. Small RNA alignments previously identified are displayed by black bars, with size shown in nt. CAAT-like elements, TATA box, asf-1 region and TSS are also displayed (boxed). TR, tandem repeat.

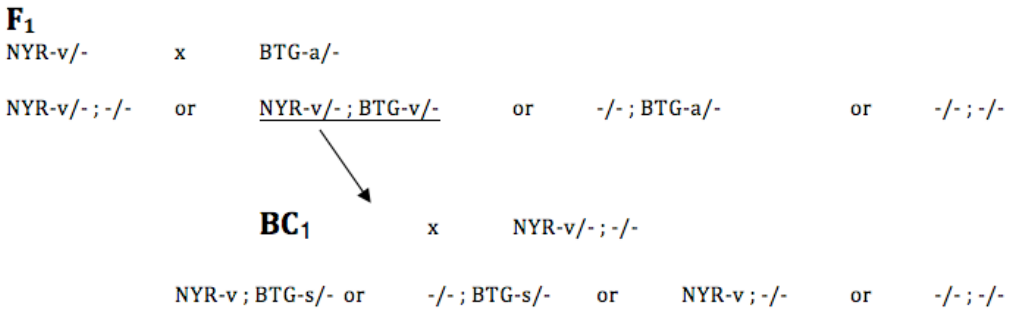


Figure 49: Crossing strategy for the genetic introgression of NYR-v and BTG-a transgenes. Hemizygous NYR-v plants were initially crossed with hemizygous BTG-a generating the F1 population. NYR-v/-; BTG-v/- plants were subsequently backcrossed with B73 WT and separately NYR generating two other populations. BC, Backcross.

absent of transgenes (Table.12).

Prior to genotyping, variegated colour was observed in the stem and leaves of a select number of plants, while others remained fully coloured (Fig.50A). Following genotyping, it became evident that all individuals showing variegated expression were NYR-v/-; BTG-v/-, all fully coloured individuals were BTG-a/- and all uncoloured plants either NYR/- or had no transgenes present. The NYR-v/-; BTG-v/- individuals selected in Fig.50A show the range of colour variegation exhibited. Collectively, these data show that NYR-v silencing can be passed in *trans* to BTG-a.

Table 12: Progeny generated from two independent crosses of NYR-v with BTG-a.

Genotype	BTG expression	Expected freq.	Observed number	Observed freq.
NYR-v/-; BTG-v/-	variegated	25%	5	25.00%
NYR-v/-; -/-	n/a	25%	4	20.00%
-/-; BTG-a/-	constitutive	25%	5	25.00%
-/-; -/-	n/a	25%	6	30.00%
		Total	20	
Genotype	BTG expression	Expected freq.	Observed number	Observed freq.
NYR-v/-; BTG-v/-	variegated	25%	4	20.00%
NYR-v/-; -/-	n/a	25%	5	25.00%
-/-; BTG-a/-	constitutive	25%	4	20.00%
-/-; -/-	n/a	25%	7	35.00%
		Total	20	

5.2.3 Methylation analysis of BTG-v epiallele

5.2.3.1 Methylation analysis by McrPCR McrPCR analysis was conducted on each phenotypic class identified (Fig.50B). Three BTG-a/- individuals all possessed hypomethylation of the BTG 35S promoter and lacked the NYR-v promoter. Three NYR-v/- individuals without colour did not possess BTG and showed hypermethylation of the NYR-v e35S promoter. Finally, three

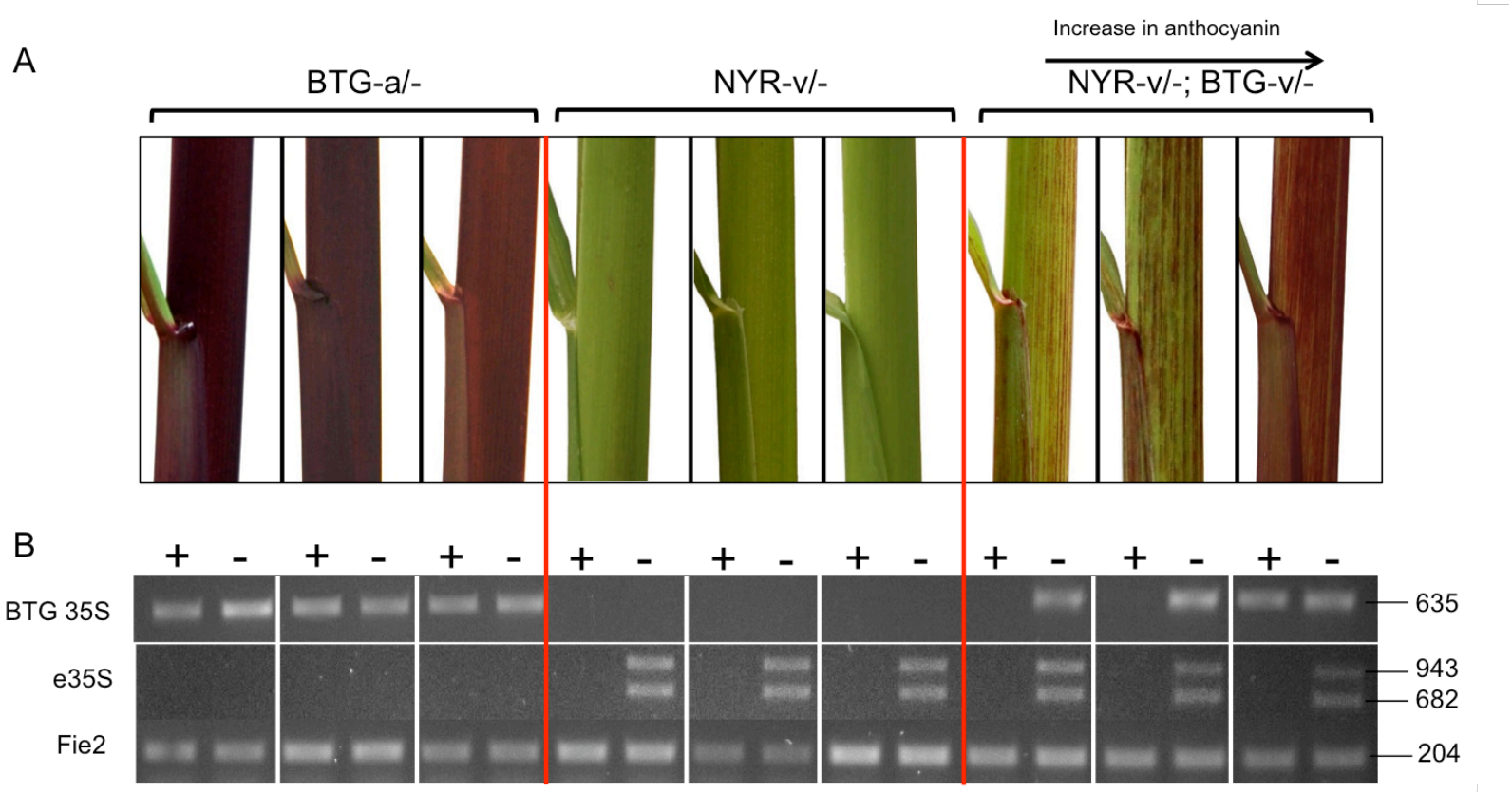


Figure 50: Phenotypic and DNA methylation analysis of the NYR-v and BTG-a transgenics. (A) Stem pigmentation of three BTG-a/-, three NYR-v/- and three NYR-v/-; BTG-v/- (of varying intensities of anthocyanin pigment production). (B) McrBC digestion and PCR amplification of of the NYR-v e35S, BTG-a/v 35S promoters (utilising a common forward primer for the promoters) and *Fie2* (unmethylated control) in each of the individuals displayed above in A. NYR-v e35S PCR produces multiple amplicons due to amplification of either one or both tandem elements. PCR product sizes (in bp) is indicated on right hand side of the panel.

NYR-v/-; BTG-v/- individuals showed hypermethylation of the BTG 35S promoter in two plants and hypomethylation in the third, which also possessed considerably more colour (still noticeably silenced). In addition, these individuals showed hypermethylation of the NYR-v e35S promoter.

These results show that NYR-v can induce silencing of BTG-a by directing *de novo* DNA methylation at the promoter.

5.2.3.2 Methylation analysis by bisulfite sequencing For a more in depth study of DNA methylation at the BTG promoter bisulfite sequencing was conducted. Hemizygous BTG-a/- and NYR-v/-; BTG-v/- plants, one presenting strong silencing of BTG and the other weak silencing of BTG (as indicated by lesser or greater amounts of colour) (Fig.51). The average methylation across the promoter region was observed as 0.8% in BTG-a/-, 54.0% in NYR-v/-; BTG-v/- (strong silencing) and 0.7% in NYR-v/-; BTG-v/- (weak silencing). In addition, NYR-v/-; BTG-v/- (strong silencing) shows a region between -404 to -323 with a complete absence of methylation in the CHH context and low amounts of methylation in CG and CHG contexts (only two CG positions with 80.0% and 0.0% methylation and two CHG positions showing 30.0% and 0.0% methylation) (Fig.52). The promoter 3' region (-87 to -11) of BTG-a/- and NYR-v/-; BTG-v/- (weak silencing) individuals showed no methylation, however NYR-v/-; BTG-v/- (strong silencing) shows an average of 56.4% methylation and 80.0% methylation in the *asf-1* region (100.0% methylation of both CG positions at positions -78 and -66) (Fig.53). Furthermore, observation of individual sites of methylation across the clones (see 76) analysed reveals no obvious tracks with reduced methylation that could produce a variegated expression. The methylation data was also displayed showing profiles of individual clones across the whole BTG 35S promoter and just at the *asf-1* element (Fig.54). These data show that classifying the methylation levels of clones find no difference in patterns between BTG-a/- and NYR-v/-; BTG-v/- (weak silencing) plants at the *asf-1* element or promoter as a whole.

Collectively, these data appear to show that the intensity of BTG silencing correlates with the amount of DNA methylation present at the promoter.

5.2.4 Stability of BTG silencing

Following interaction of the NYR-v and BTG-a, I crossed NYR-v/-; BTG-v/- plants by NYR-v/- segregating the transgenes and for other progeny maintaining another generation with NYR-v (crossing strategy shown in Fig.49). These crosses included both the NYR-v/-; BTG-v/- strong and weak silencing plants previously analysed. Primarily this allows for the observation of BTG

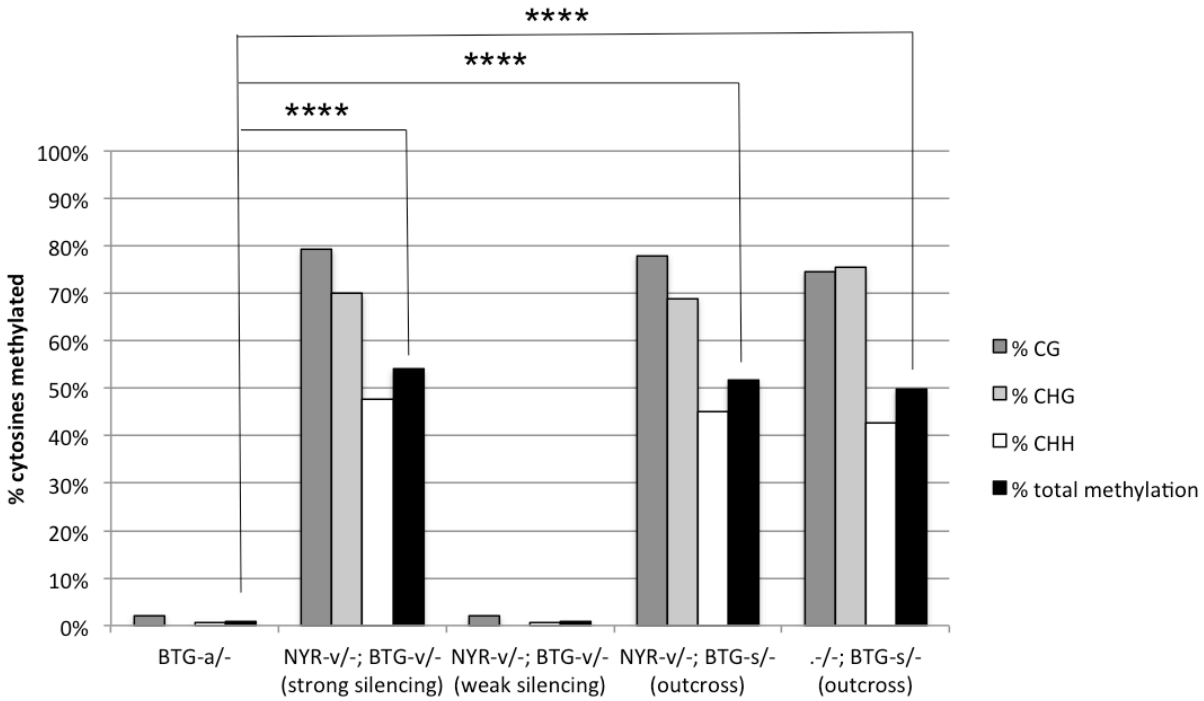


Figure 51: Bisulfite analysis of the BTG promoter with and without NYR-v. PCR amplification and sequencing of bisulfite treated gDNA focusing on the BTG 35S promoter region of BTG-a/-, NYR-v/-; BTG-v/- (strong and weak silencing), NYR-v/-; BTG-s/- (outcrossed) and -/-; BTG-s/- (outcrossed) plants. The level of methylation is reported as the average percentage of total cytosines exhibiting methylation established from 10 independent clones for each data set from position -420 to +2 of BTG transgene, a total of 107 cytosines. Asterisks denote significant difference compared with BTG-a/- of total methylation ($p<0.0001$) as established by Wilcoxon two sample rank test. Raw data presented in Appendix Table.26 and Lollipop diagram of methylation state of individual cytosine positions in all clones – Appendix Fig.76.

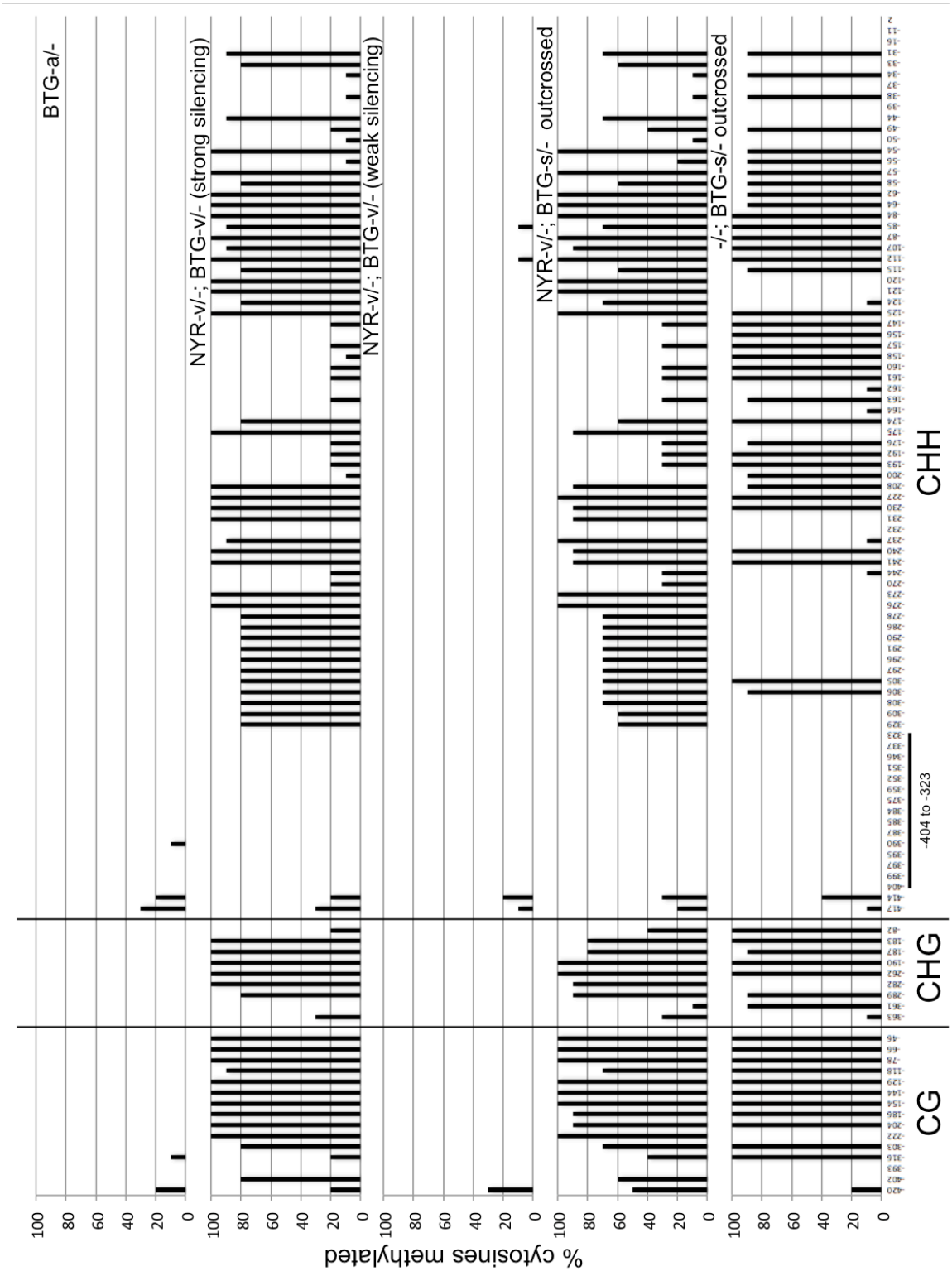


Figure 52: Methylation analysis of BTG 35S promoter showing CG, CHG and CHH methylation with and without NYR-v introgression. Bisulfite analysis of DNA methylation in the BTG promoter of BTG-a/-, NYR-v/-; BTG-v/- (strong and weak silencing), NYR-s/-; BTG-v/- (outcrossed) and -/-; BTG-s/- (outcrossed) plants. The level of methylation is reported as the percentage of individual cytosine locations exhibiting methylation, established from 10 independent clones, spanning -420 to +2 of BTG (cytosine positions marked on the x-axis and have been split between CG, CHG and CHH contexts), a total of 107 cytosines. Positions -404 to -323 (underlined) represent regions of conserved reduction in methylation.

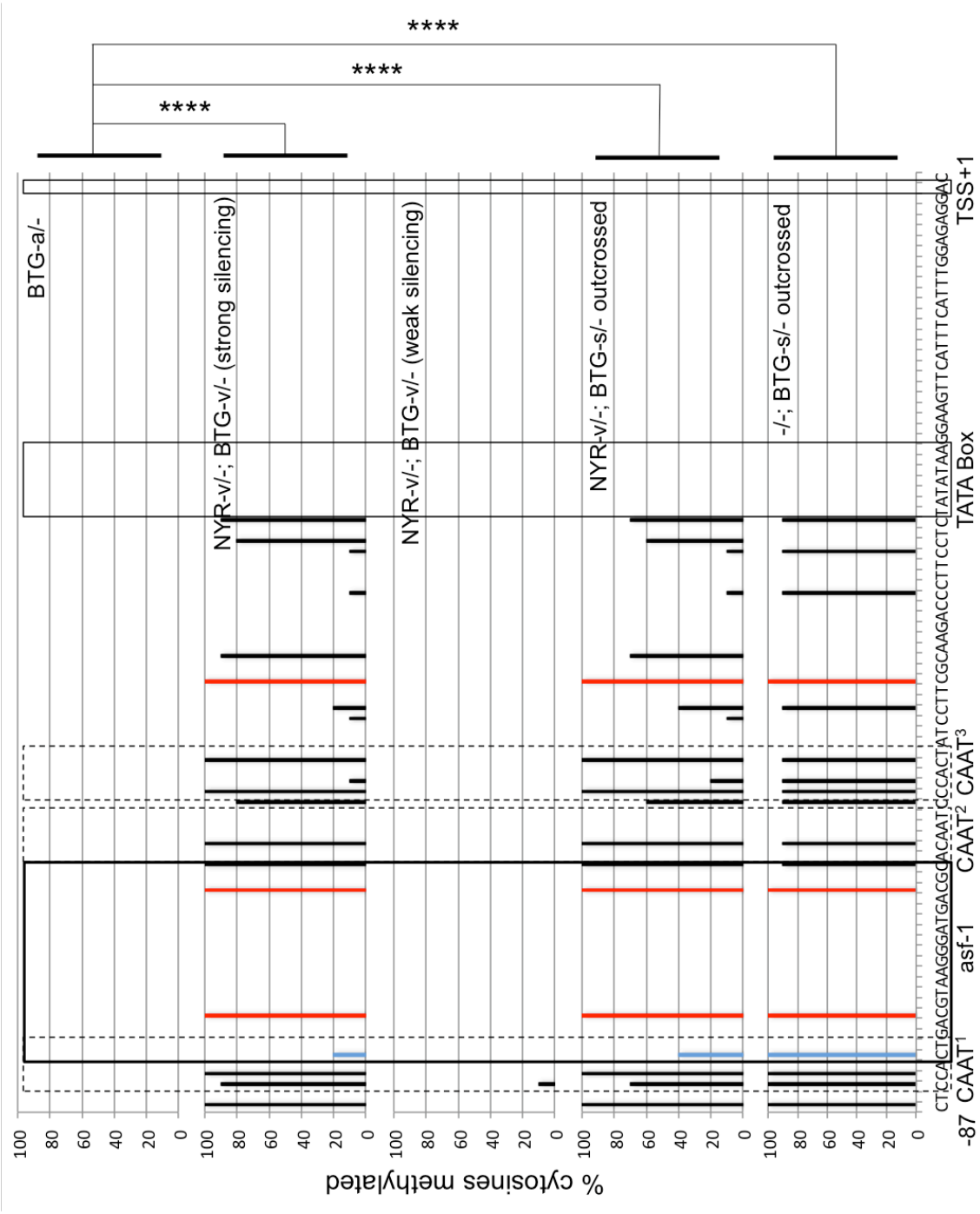


Figure 53: Methylation profile of BTG unique 3' promoter region with and without NYR-v introgression. Methylation levels of cytosine positions between -87 and +1 of the BTG 35S promoter, displaying CAAT-like elements, asf-1 element, TATA box and TSS (+1) in BTG-a/-, NYR-v/-; BTG-v/- (strong and weak silencing), NYR-v/-; BTG-s/- (outcrossed) and -/-; BTG-s/- (outcrossed) plants. The level of methylation is reported as the percentage of individual cytosine locations exhibiting methylation, established from 10 independent clones (position number marked on the x-axis). The asterisks denote the significant difference between average levels of methylation compared with BTG-a/- ($p<0.0001$) of the data sets as established by Wilcoxon two sample rank test.

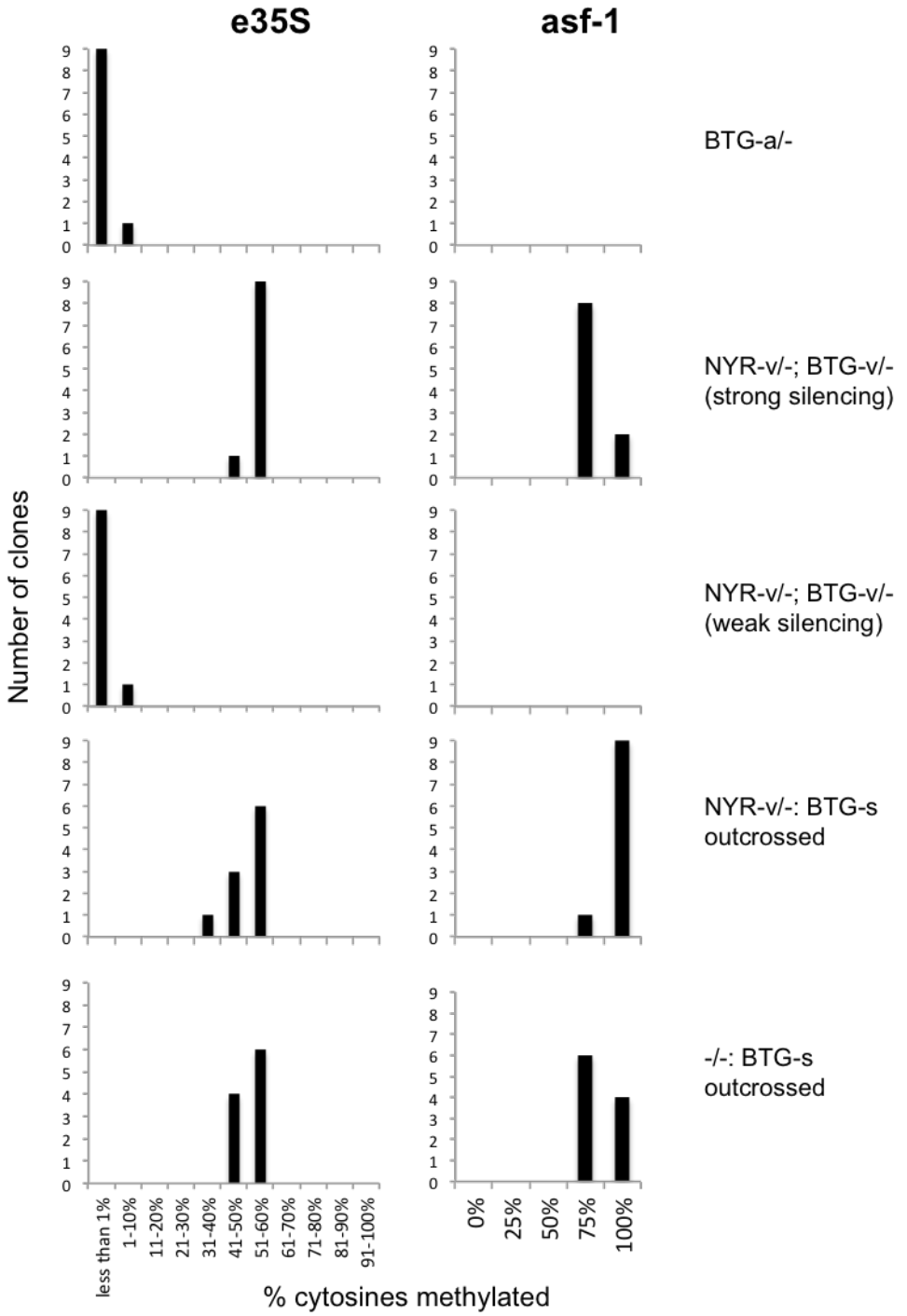


Figure 54: DNA methylation status of individual clones across the BTG 35S promoter with and without NYR-v introgression. The level of methylation for individual clones is reported as a percentage of cytosines exhibiting methylation in the BTG 35S promoter and the asf-1 region of the promoter (only four cytosines) in BTG-a/-, NYR-v/-: BTG-v/- (strong and weak silencing), NYR-v/-; BTG-s/- (outcrossed) and -/-; BTG-s/- (outcrossed) plants. 10 clones had been sequenced for each dataset.

reporter expression post-NYR-v. Five independent crosses were conducted and 20 progeny plants were then grown and assessed for BTG colour expression in the stem and leaves. All plants grown showed no pigmentation, the plants from a single line were genotyped showing eight NYR-v; -/-, two -/-; BTG-s/-, seven NYR-v/-; BTG-s/- and three -/-; -/- (Appendix Table.28). All plants observed from the outcrossed generation showed no colour expression and NYR-v expression appeared unaffected (Fig.55A).

These data show that silencing of BTG has intensified in its second generation, with and without the presence of NYR-v.

5.2.5 Methylation analysis of stable BTG-s plants

5.2.5.1 Methylation analysis by McrPCR The BTG and NYR promoters of two individuals from each genotype were analysed by McrPCR (Fig.55B). NYR-v/-; -/- progeny showed hypermethylation of the NYR-v promoter, -/-; BTG-s/- showed hypermethylation of the BTG promoter, NYR-v/-; BTG-s/- show hypermethylation of both the BTG-s and NYR-v promoters and BTG-a/- (not crossed with NYR-v) shows hypomethylation of the BTG-a promoter.

5.2.5.2 Methylation analysis by bisulfite sequencing Subsequent analysis by bisulfite sequencing of the outcrossed individuals was conducted on the BTG promoter region (Fig.51). These results reveal an average methylation of 51.7% and 49.9% in NYR-v/-; BTG-s/- and -/-; BTG-s/- respectively (compared with 54.0% in NYR-v/-; BTG-v/- (strong silencing) after the first cross with NYR-v). Moreover, both share an absence of CHH methylation between positions -404 and -323 (Fig.52), previously seen in NYR-v/-; BTG-v/- following the first crossing (see 5.2.3).

Analysis of the promoter 3' region shows an average methylation of 54.4% and 64% of outcrossed plants NYR-v/-; BTG-s/- and -/-; BTG-s/- respectively (Fig.53). The *asf-1* region shows average methylation of 97.5% (CG positions -78 and -66, both at 100.0%) and 85.0% (Both at 100.0%) of NYR-v/-; BTG-s/- and -/-; BTG-s/- plants respectively. In addition, the CAAT-like elements show considerable methylation (Appendix Table.26). The NYR-v/-; BTG-v/- (strong silencing) generated from the first crossing as previously indicated shows 56.4% methylation in the promoter 3' region, with 80.0% methylation in the *asf-1* region (100.0% at both CG positions).

The methylation profiles of individual clones used for bisulfite sequencing of the BTG promoter and *asf-1* element was also displayed (Fig.54). This shows methylation profiles across the whole promoter are very similar between BTG-s plants (both outcrossed individuals) and the previous generations NYR-v/-; BTG-v/- (strong silencing) individual. A higher number of clones are

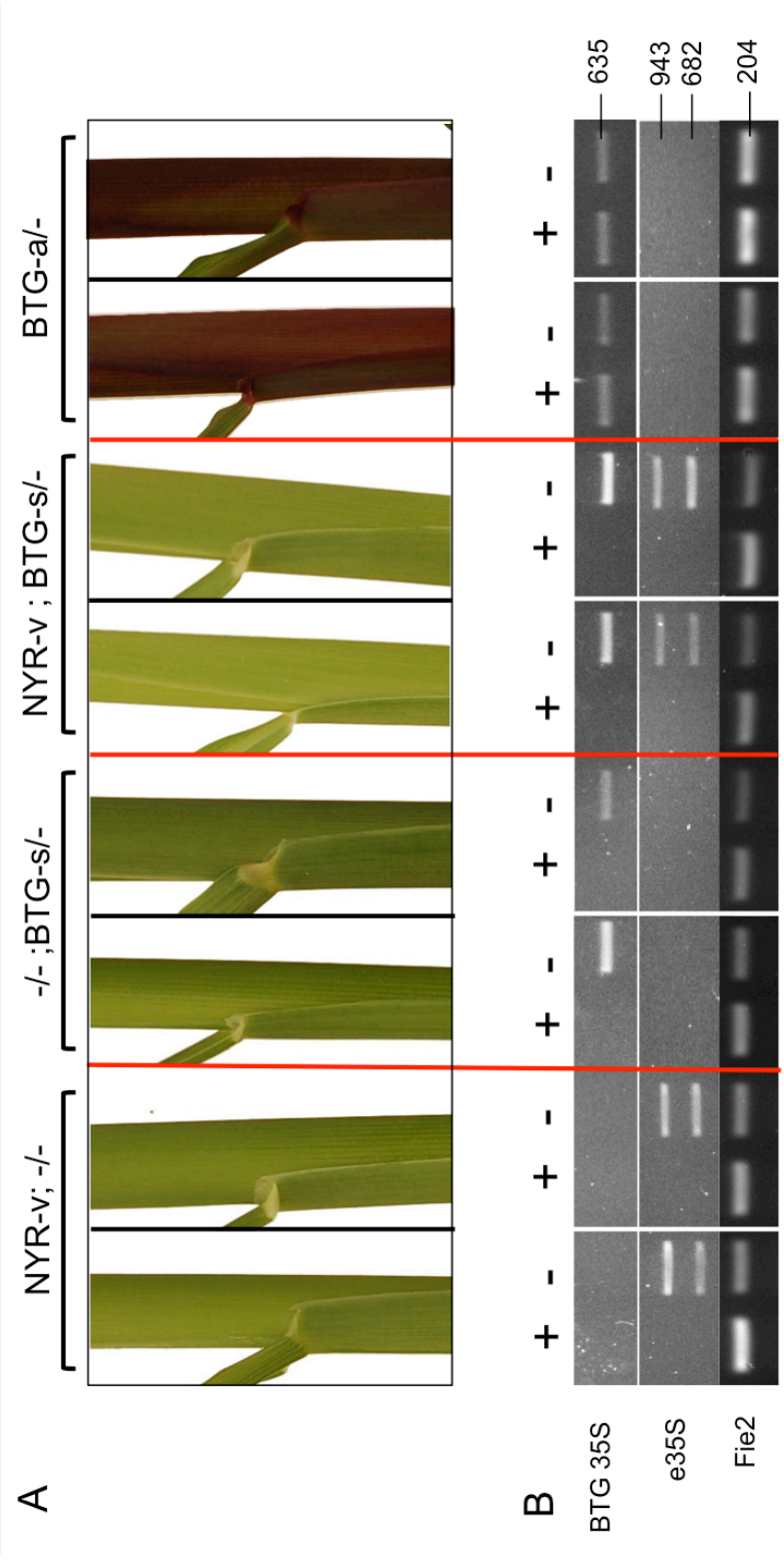


Figure 55: Phenotypic and DNA methylation analysis of the NYR-v and BTG transgenics following outcrossing. (A) Stem pigmentation of progeny of NYR-v/-; BTG-v/- plants crossed with NYR-v/-, two NYR-v; -/-, two -/-; BTG-s/-, two NYR-v; BTG-s/- and two individuals BTG-a/- that have not been in contact with NYR-v. (B) McrBC digestion and amplification of BTG 35S, NYR-v e35S promoters and *Fie2* (unmethylated control). PCR product sizes (in bp) is indicated on right hand side of the panel.

observed with 100.0% methylation of the cytosine positions within the *asf-1* region. NYR-v/-; BTG-v/- has 2 clones (8 clones 75% methylation), NYR-v/-; BTG-s/- (outcrossed) has 9 clones (1 clone 75% methylated) and -/-; BTG-s/- (outcrossed) has 4 clones (6 clones 75% methylated) with 100.0% methylation of the *asf-1* element. Of the 8 clones showing 75% methylation in NYR-v/-; BTG-v/- (strong silencing) individual, the lack of methylation at a single asymmetric (CHH) location was found to be responsible (position -82).

These data show that the *de novo* methylation deposited at the BTG-s promoter is meiotically stable.

5.2.6 Paramutation-like effects in maize transgenes

The epigenetic *trans*-silencing observed between NYR-v and BTG-a indicates the possibility of paramutation-like effects between transgenes. To test this hypothesis NYR-v was crossed with a third transgene, Cell wall YFP Reporter (CYR - referred to as CYR-a (CYR-active) (Appendix Fig.78)), which utilizes an identical tandem e35S promoter present in NYR-v to demonstrate its ability to silence in *trans* (Fig.56). Following the outcrossing of BTG-v, (which generated BTG-s individuals) progeny can then be crossed with CYR-a, observing expression of the CYR transgene. Silencing of CYR-a dictates that BTG-v was paramutagenic, and able to silence in *trans*, without the original silencing element, NYR-v. Paramutation is also known to be heavily associated with DNA methylation, and components of the RdDM pathway are essential. Therefore, the DNA methylation and 24 nt siRNAs detected at the NYR-v promoter, the only region of homology between the transgenes, suggests that paramutation is possible.

5.2.7 Phenotypic analysis of interaction between NYR-v and CYR-a

Crosses were carried out between hemizygous NYR-v and CYR-a transgenics, 30 plants were then germinated and grown from three independent crosses (Appendix Table.29). Confocal microscopic analysis was conducted in roots and leaf samples of progeny plants with CYR. This analysis revealed that NYR-v can induce a variegated expression in CYR-a (Fig.57).

Closer analysis was also conducted on the spatial distribution of NYR and CYR active cells. Confocal microscopy was used to to identify nuclear and/or cell wall expressing cells (NYR-a and CYR-a at the single cell level). An average of 7.2%, 14.3%, 20.4% and 19.4% of CYR-a expressing cells also expressed the NYR transgene in four plants respectively (Fig.58). In all cases, less cells were observed expressing NYR, than CYR.

Collectively, these data indicate that NYR-v efficiently directs silencing of CYR-a.

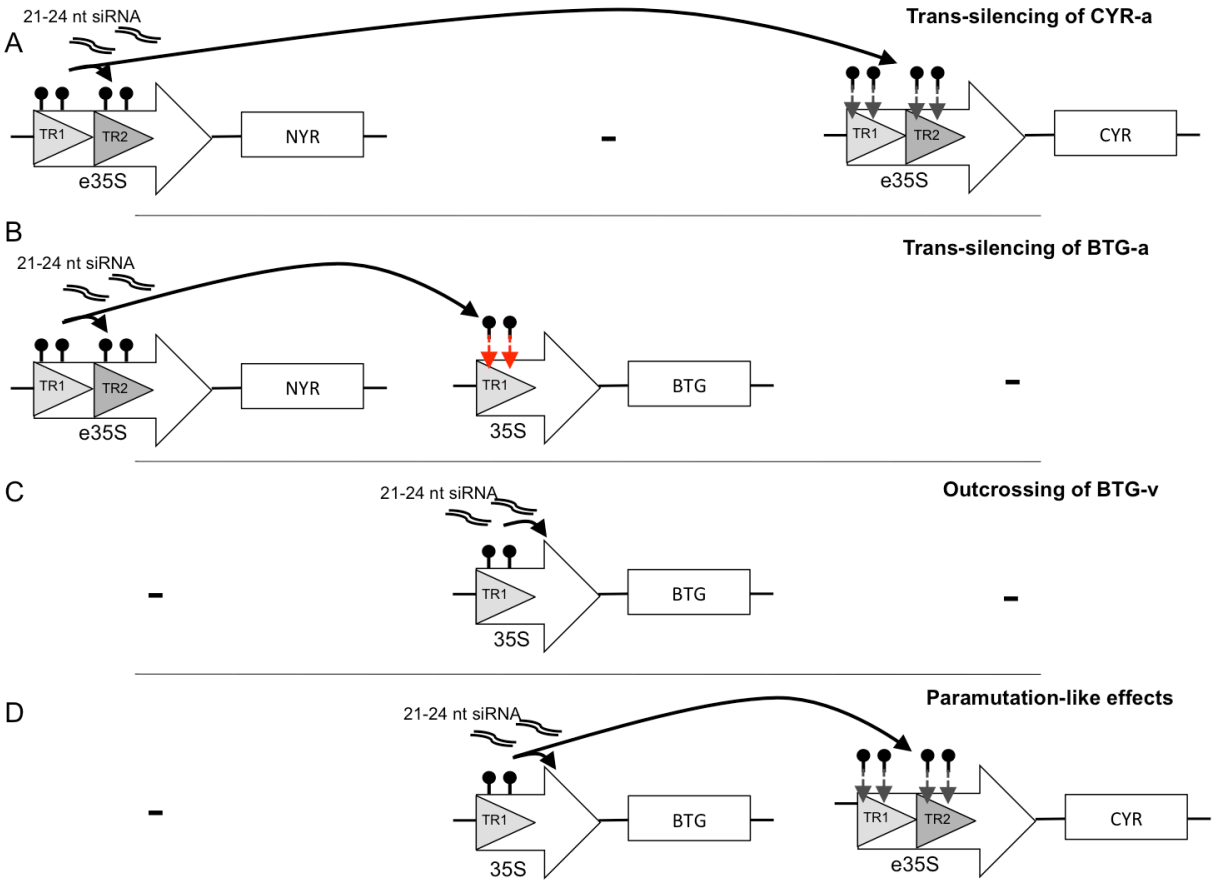


Figure 56: Schematic diagram of hypothesised mode of NYR-v paramutation. Possible route of paramutation, silencing transgenes over three consecutive generations (B-D). (A) NYR-v/-; CYR-v/- plants (NYR-v known to *trans*-silence CYR-a). (B) NYR-v/-; BTG-v/- plants (NYR-v known to *trans*-silence BTG-a). (C) -/-; BTG-s/- plants (outcrossing BTG, removing NYR, shown to silence BTG, methylation remains. (D) BTG-s/-; CYR/- plants (hypothesised BTG *trans*-silencing of CYR). Black lollipops represent methylation, red arrows represent action of *de novo* methylation, grey arrows represent hypothesised *de novo* methylation and black arrows the activity/hypothesised activity of 21-24 nt siRNAs.

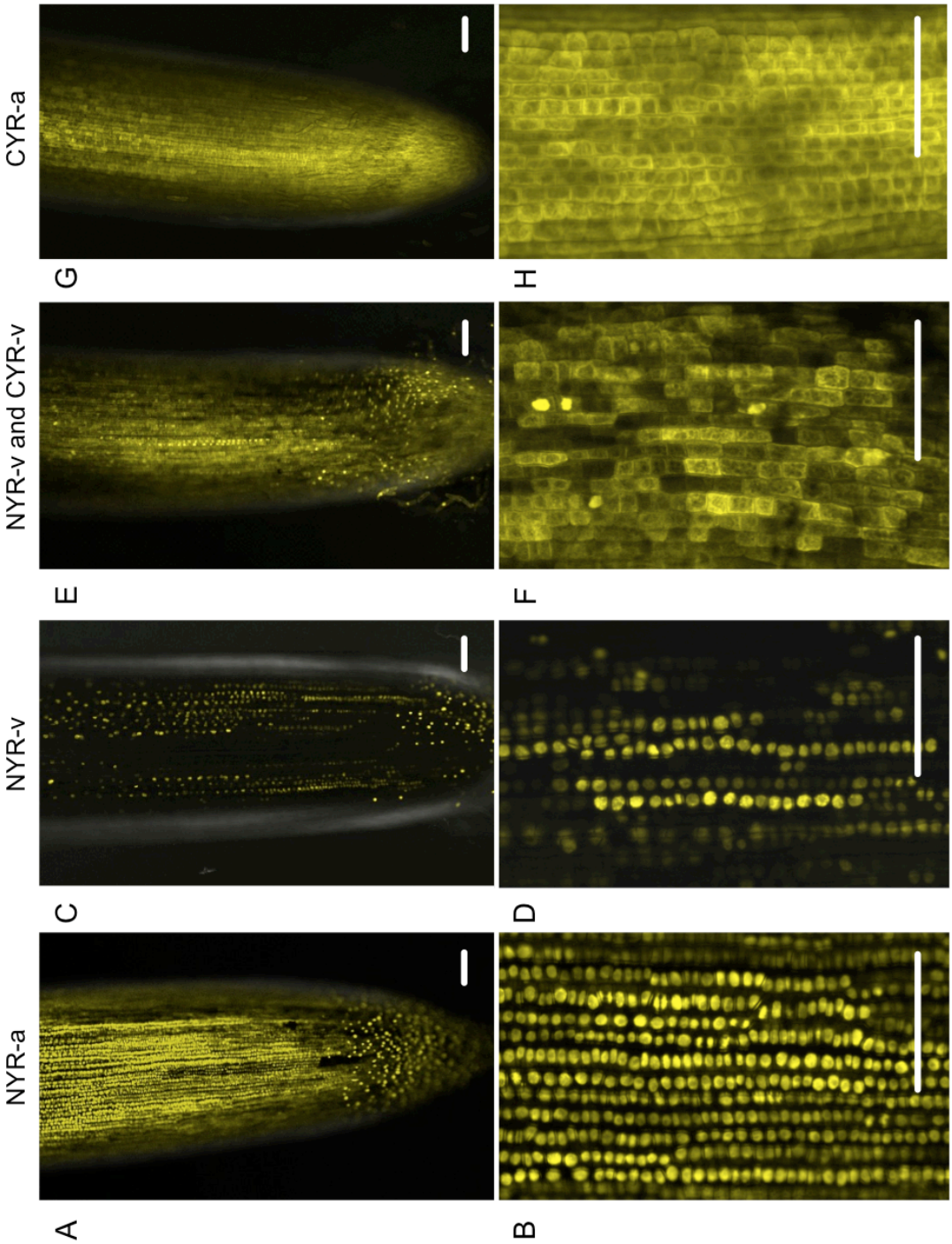


Figure 57: Confocal microscopic analysis of NYR-v/a and CYR-v/a in roots. YFP imaging of (A,B) NYR-a reporter expression, (C,D) NYR-v reporter expression, (E,F) NYR-v and CYR-v reporter expression and (G,H) CYR-a expression in root tissue. Scale bar = 100 µm

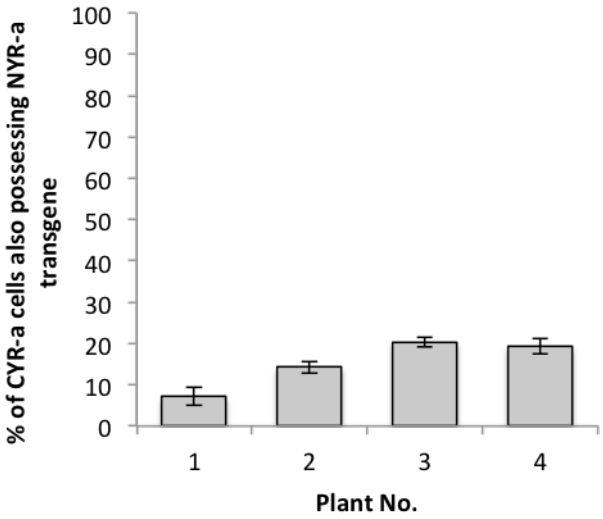


Figure 58: Percentage of CYR-a expressing cells that also show NYR-a expression. Average percentage of observed cells counted showing CYR-a expression from NYR-v/-; CYR-v/- plants that also show NYR-a expression, as detected by confocal microscopy. Roots were tested from four plants (labeled 1-4) and average percentages plotted. Error bars show the range of results between four locations imaged at each root prior to averaging.

5.2.8 Phenotypic analysis of BTG-s and CYR-a progeny

NYR-v has now been shown to induce silencing in both BTG-a and CYR-a. BTG-s/- (outcrossed from NYR-v) plants were crossed with CYR-a/- in three independent cross events. 27 plants were grown from each cross event and CYR expression was assessed by confocal microscopy. However, only constitutive expression of YFP was observed in the roots and leaves of CYR-a/-; BTG-s/- genotypes in progeny generated from three independent crosses (Table.13). However, an additional 40 plants were later grown from two independent crosses, to check if silenced CYR expression was a rarer event. Of the 80 plants grown, 7 were identified exhibting a variegated expression (CYR-v/-; BTG-s/-). Furthermore, BTG-a/- crosses with CYR-a/- were previously shown to have no effect on BTG or CYR expression (data not shown).

These data indicate that BTG-s can direct silencing at CYR-a

Table 13: Progeny generated from three independent crosses of CYR-a and BTG-s.

Genotype	CYR expression	Expected freq.	Observed number	Observed freq.
CYR-a/-; BTG-s/-	constitutive	25%	5	18.52%
CYR-a/-; -/-	constitutive	25%	9	33.33%
-/-; BTG-s/-	n/a	25%	6	22.22%
-/-; -/-	n/a	25%	7	25.93%
		Total	27	
Genotype	CYR expression	Expected freq.	Observed number	Observed freq.
CYR-a/-; BTG-s/-	constitutive	25%	3	11.11%
CYR-a/-; -/-	constitutive	25%	8	29.63%
-/-; BTG-s/-	n/a	25%	5	18.52%
-/-; -/-	n/a	25%	11	40.74%
		Total	27	
Genotype	CYR expression	Expected freq.	Observed number	Observed freq.
CYR-a/-; BTG-s/-	constitutive	25%	7	25.93%
CYR-a/-; -/-	constitutive	25%	8	29.63%
-/-; BTG-s/-	n/a	25%	6	22.22%
-/-; -/-	n/a	25%	6	22.22%
		Total	27	

5.2.9 Assessing epigenetic communication between seeds components

The *trans*-silencing interaction of NYR-v and BTG-a provides a tool to study epigenetic communication via small RNA movement during seed development. Thus, I designed experiments in order to study the possibility of the transfer of epigenetic states between the endosperm and embryo of developing maize seeds. It has been theorised that small RNAs move from the endosperm to the embryo, in order to silence TEs and other unfavourable regions (Gehring et al. 2009; Mosher and Melnyk 2010). This report has already evidenced small RNAs (21 and 24 nt siRNAs) in the direction of DNA methylation and initiating silencing of NYR-v and also its ability to silence BTG in *trans* due to promoter sequence homology. Therefore a tissue containing a source of small RNAs (NYR-v) and a neighbouring tissue containing a target for the small RNAs (BTG-a) could hypothetically communicate epigenetically. Two techniques were employed to test this hypothesis, first bringing together genetically dissimilar endosperm and embryos by heterofertilisation, and secondly by grafting.

5.2.10 Heterofertilisation events in BTG-a by NYR-v crosses

Heterofertilisation is a phenomenon known to occur in maize, whereby a small percentage of seeds formed are genetically dissimilar (Gao et al. 2011). Seeds from seven independent crosses of BTG-a/- by NYR-v/- allowed for successful germination and growth of 873 seeds, grown in two experiments due to space constraints (210 in expt. no.1, 663 in expt. no.2) (number and

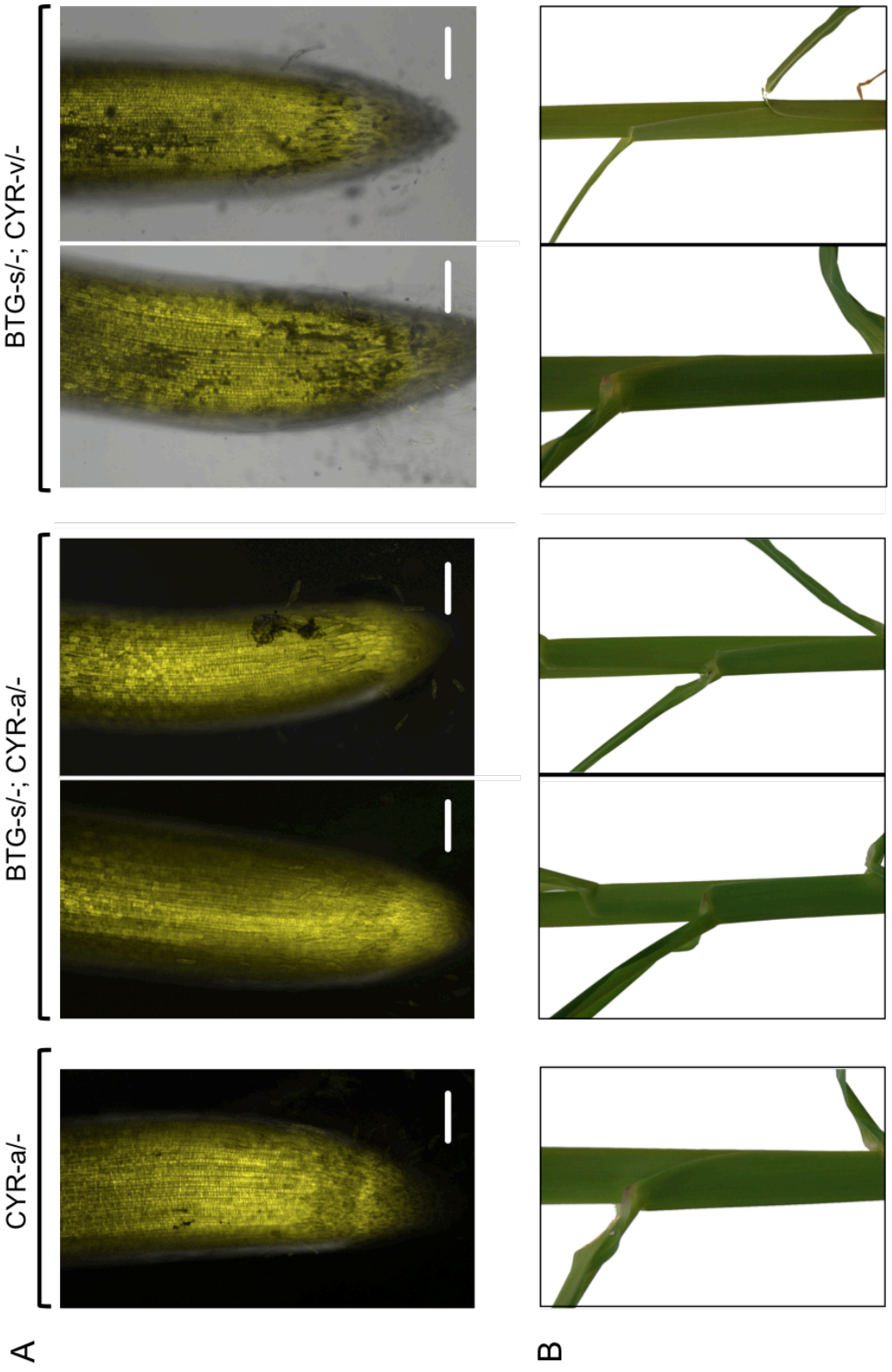


Figure 59: Confocal microscopic and phenotypic analysis of CYR and BTG expression following introgression. (A) Confocal microscopic imaging of CYR expression in roots from a CYR-a/- individual, two images from CYR-a/-; BTG-s/- individuals and two images from CYR-v/-; BTG-s/- individuals. Scale bar = 200 µm. (B) Stem images of the same plants, showing BTG-s phenotype (no colour).

frequencies shown in Table.14). This total would therefore allow for between 4.4 and 43.7 heterofertilisation events occurring, assuming that 0.5-5% of progeny are affected (Wu et al. 2013) (For more information on heterofertilisation see 5.1.5). Phenotypic analysis of seeds and ears prior to growth showing different potencies of colour expression in the aleurone. A subset of four ears displaying differing proportions of coloured seed possess 0.0%, 5.8%, 27.1% and 44.9% of seeds with colour (Fig.61). The maximum expected percentage of which would be 50.0% based on the cross conducted. Furthermore, the variation in colour potency is recorded as strongly or weakly pigmented (Fig.61B). The weak pigmentation could be a direct result of the silencing action of NYR-v.

Phenotypic analysis of stem and leaves of the 663 plants in expt. no.2 again reveals varying levels of variegation and plants were classified into one of five categories dependent on the amount of colour expression present (Fig.60). The first category representing strong silencing of BTG (less colour) and category five as weak silencing (more colour). Cat. 1 showed 3 individuals (1.8%), cat. 2 showed 35 (21.5%), cat.3 showed 75 (46.0%), cat. 4 showed 47 (28.8%) and cat 5. showed 4 (2.5%).

Plants with full colour or variegation were tested for the presence of NYR-v. 231 plants were NYR-v/-; BTG-v/-, all showing colour variegation and 236 plants were -/-; BTG-a/-, all showing full colour expression between independent experiments. In addition, all uncoloured individuals did not possess BTG.

However, no plants were identified that presented a variegated colour phenotype without NYR-v, indicating that heterofertilisation frequency is lower than reported or that silencing has not efficiently transmitted from endosperm to embryo.

Table 14: Progeny derived from hemizygous NYR-v and BTG-a crosses. Experiment 1 is the total number from three independent cross events and Experiment 2 from an additional four independent cross events.

				Observed	
Expt. No.	Genotype	Colour phenotype	Expected Frequency	Number	Frequency
1	NYR-v/-; BTG-v/-	variegated	25%	68	32.40%
	-/-; BTG-a/-	full colour	25%	59	28.10%
	NYR-v/-; -/- or -/-: -/-	n/a	50%	83	39.52%
				210	
2	NYR-v/-; BTG-v/-	variegated	25%	163	24.60%
	-/-; BTG-a/-	full colour	25%	177	26.70%
	NYR-v/-; -/- or -/-: -/-	n/a	50%	323	48.725
				663	

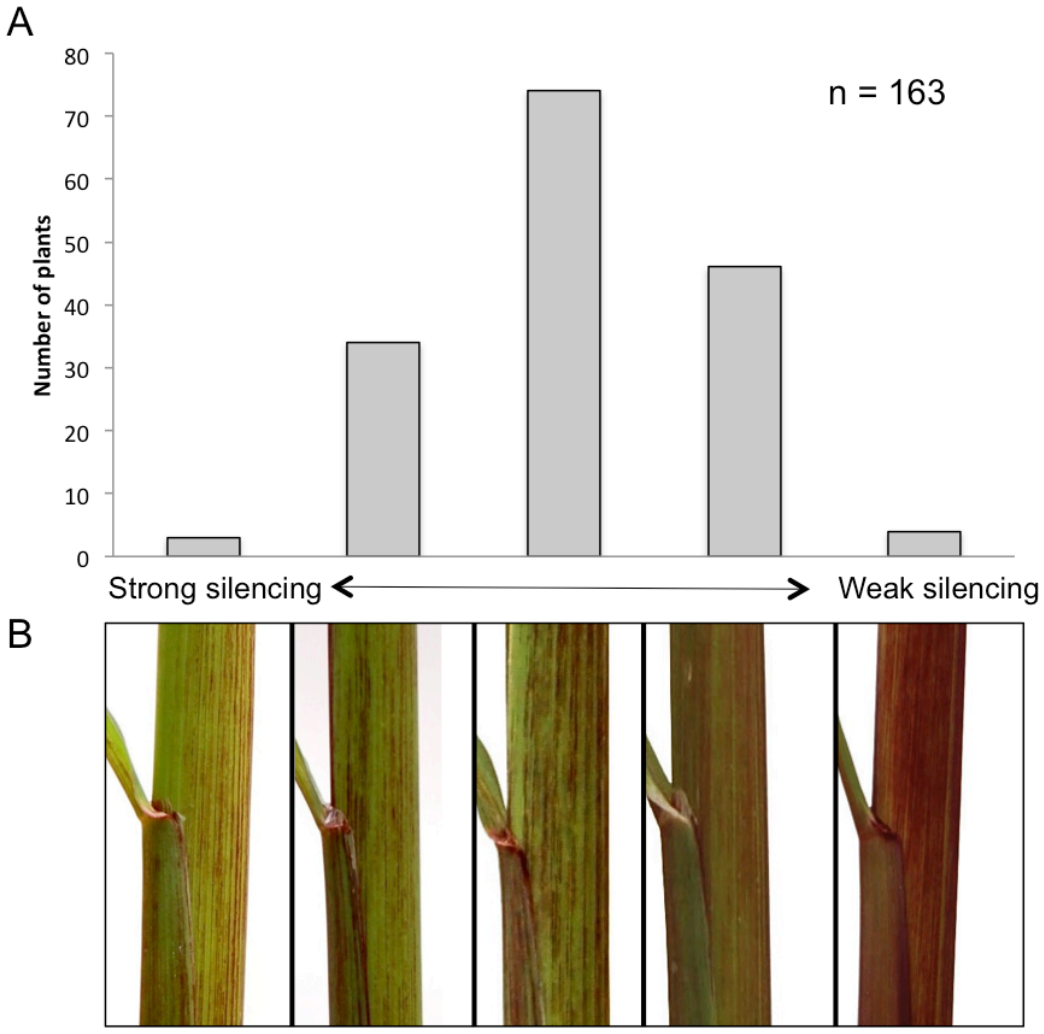


Figure 60: Distribution of the potency of observed BTG-v silencing by NYR-v. (A) Classification of 163 NYR-v/-; BTG-v/- plants into five categories of silencing potency, based on visual assessment of abundance of colour. (B) Stem images of BTG-v individuals representing the five categories.

A

B

C

0	5	17	9	Strongly pigmented seeds
0	7	65	79	Weakly pigmented seeds
190	204	303	196	Total seeds
0.0%	5.8%	27.1%	44.9%	Seeds pigmented

Figure 61: Phenotypic analysis of observed seeds resulting from NYR-v and BTG-a crosses. (A) Images of four ears derived from BTG-a/- by NYR-v/- crosses showing variation in the abundance and intensity of colour pigmentation of individual seeds (B) Images of a subset of seeds from the above ears. (C) Table showing number of seeds and pigment levels from each ear. Scale bar = 1 cm

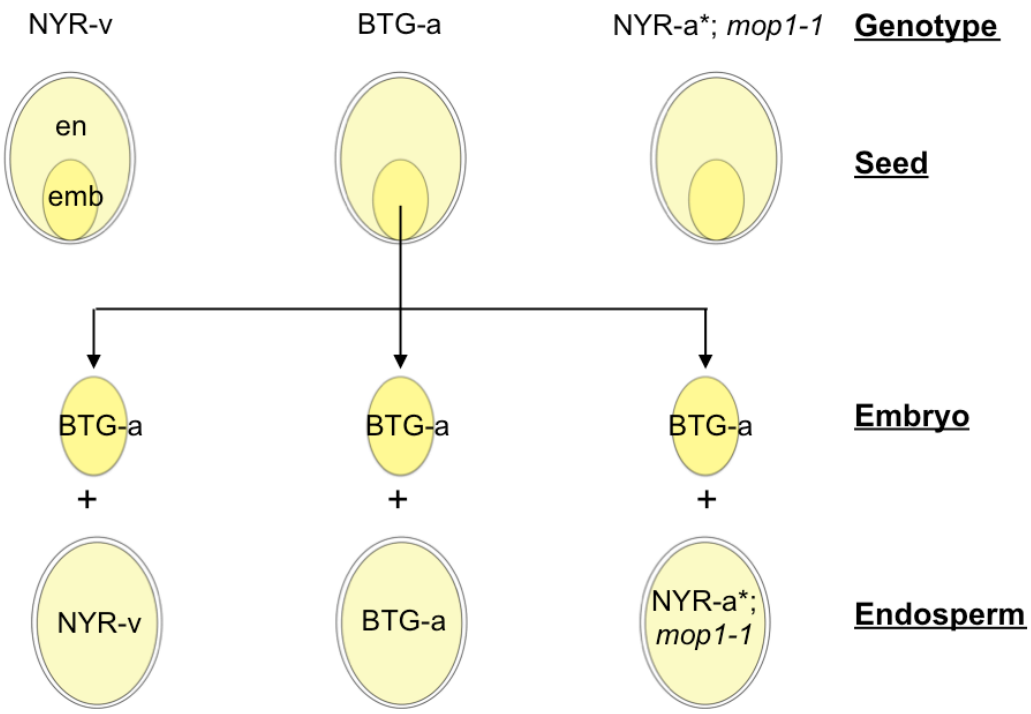


Figure 62: Strategy for endosperm and embryo grafting. Schematic diagram showing crosses performed in generating seeds with the endosperm and embryo genotypes desired to produce the combinations of grafted tissues shown underneath. Only the desired genotypes are shown, and not other genotypes generated by these crosses. Light yellow ovals, endosperm; Yellow ovals, embryo.

5.2.11 Grafting endosperm and embryos

A second experiment was conducted, where grafting genetically dissimilar endosperms and embryos was used to generate mature plants. Endosperm and embryos were derived from NYR-v, BTG-a and NYR-a* *mop1-1* sibling crosses and organised as shown in the design schematic (Fig.62). I chose to use 6 DAP embryos (Fig.63A), securing embryos at the pretransitional/transitional stage in development where the meristematic region is still developing. This also maximises the time of exposure to endosperm tissue prior to germination for the movement of small RNAs. Endosperm tissue was harvested later at 10-12 DAP allowing for enough tissue to be used as a base for embryos (Fig.63C).

The number of endosperm and embryo combinations initially placed and the number of mature plants generated, which includes 15 derived from BTG-a endosperm and embryos, 41 derived from NYR-v endosperms and BTG embryos and 11 derived from NYR-a* *mop1-1* endosperm and BTG-a embryos as shown in Table.15. The number of grafted plants that were moved to upright containers is a lot lower than the initial number of endosperm and embryo grafts due to constraints on the number of containers avaliable for use.

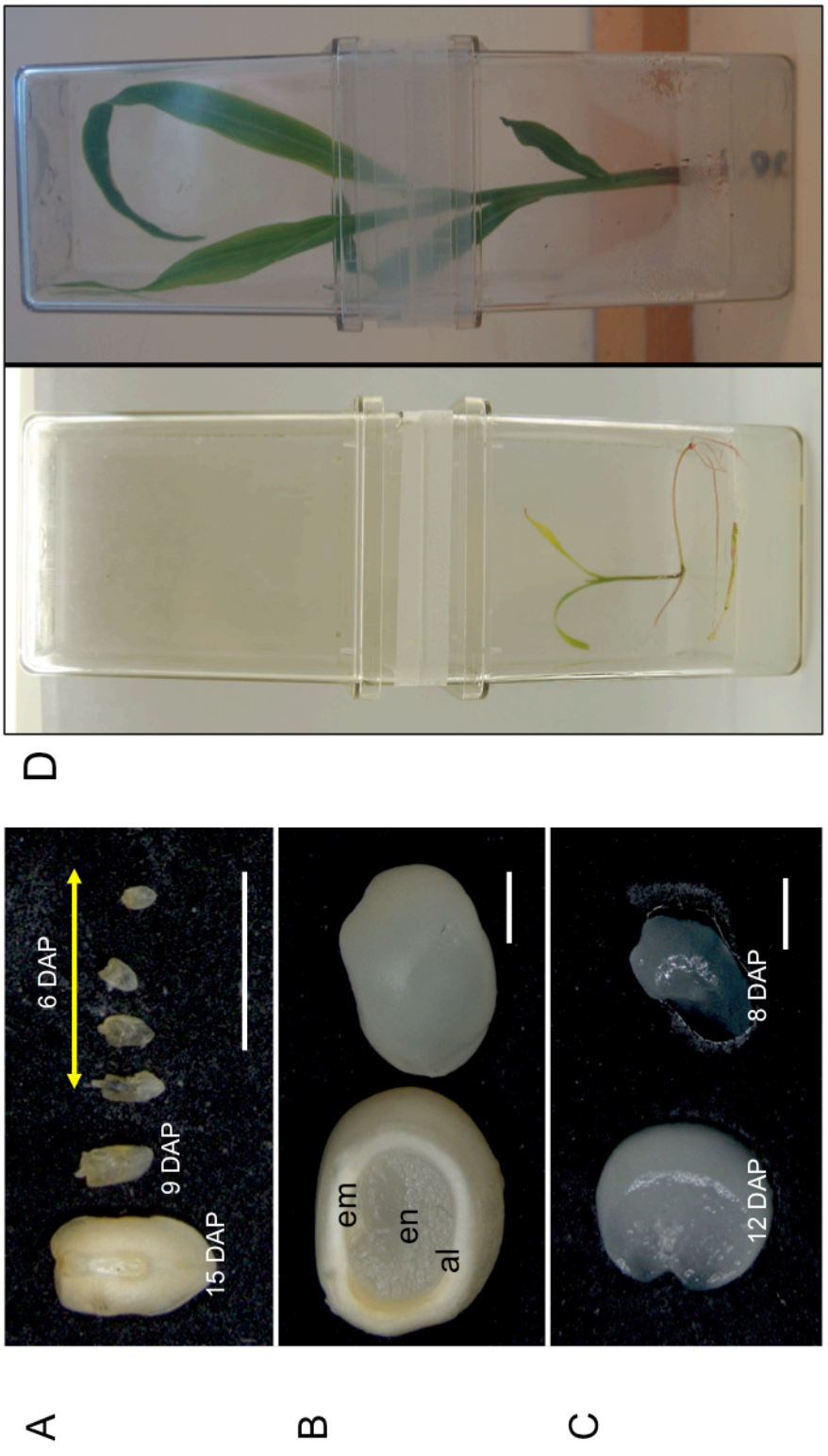


Figure 63: Components of endosperm and embryo grafting experiments. Images of (A) excised embryos: 6, 9 and 15 DAP, (B) 12 DAP seed following transversal cut, exposing endosperm and embryo, and separated endosperm, (C) 12 and 8 DAP separated endosperm. (D) Transfer of growing plantlet from endosperm/embryo graft to upright container, displaying 2 weeks and 5 weeks growth. al, aleurone; en, endosperm; em, embryo. Scale Bars = 2 mm.

Table 15: Numbers of endosperm and embryo grafting events carried out and mature plants generated.

Graft (endo. and emb.)	Total grafted	Transfered to upright container	Transfered to soil	Total survived	Unpigmented	Pigmented	
						Full colour	Variegated
BTG-a/- and BTG-a/-	95	20	16	15	9	6	0
NYR-v/- and BTG-a/-	351	68	59	41	14	27	2
NYR-a*/-; *mop1-1* and BTG-a/-	144	21	12	11	5	6	0

5.2.12 Phenotypic analysis of adult grafted plants

The stems and leaves of plants were assessed for variegation in colour expression, indicative of silencing of BTG. Two BTG plants (recorded as plant G6 and G8) generated from germination with NYR-v endosperms (out of 41 total) showed variegated colour expression in stems and sectored expression in leaves (Table.15). All other plants observed in this group and those derived from growth with BTG-a or NYR-a* *mop1-1* endosperms showed full colour expression (Fig.64A).

5.2.13 Methylation analysis of grafted plants

5.2.13.1 Methylation analysis by McrPCR McrPCR analysis was conducted, observing hypomethylation in the BTG promoter of BTG-v plants (G6 and G8) derived from growth with NYR-v endosperm, BTG-a plants grown with BTG-a endosperm and BTG-a plants grown with NYR-a*; mop1 endosperm (Fig.64B).

5.2.13.2 Methylation analysis by bisulfite sequencing Bisulfite sequencing was performed on a BTG-a individual control which had been germinated with a BTG-a/- sibling cross derived endosperm as well as the two grafted BTG-v individuals (G6 and G8) germinated with NYR-v sibling cross derived endosperms (Fig.65). Methylation levels at the BTG promoter are very low in all three individuals, showing total average methylation profiles of 0.9%, 1.1% and 1.3% for

Table 16: Progeny derived from crosses utilising grafted plants with reduced colour.

BTG-v/- (G8) x B73 WT

Genotype	BTG expression	Expected freq.	Observed number	Observed freq.
BTG/-	variegated	50%	16	53.30%
-/-	n/a	50%	14	46.70%
		Total	30	

BTG-a/- x BTG-v/- (G6)

Genotype	BTG expression	Expected freq.	Observed number	Observed freq.
BTG/- or BTG/BTG	variegated or full	75%	23	76.70%
-/-	n/a	25%	7	23.30%
		Total	30	

BTG-a control, BTG-v (G6) and BTG-v (G8) respectively. BTG-v G6 and G8 plants, displayed CHH methylation within the 3' unique region of the promoter, with 10.0% methylation detected at positions -66 (within the CAAT^2-like element) and -44 of plant G6 and 10.0% methylation detected at position -107 of plant G8 ((Fig.65) (Dissected results shown in Appendix table.27). These data indicate that DNA methylation is not present at the promoter in grafted plants showing variegated expression.

5.2.14 Phenotypic analysis of grafted BTG-v following growth of another generation

The G6 and G8 grafted individuals were then crossed by a full colour expressing grafted individual and B73 WT respectively. These crosses were performed to observe the BTG expression in the next generation and also to observe the effect that BTG-v may have upon BTG-a plants. Seeds from both cross events were germinated, 30 of which were genotyped for the presence of BTG (Table.16). All plants derived from the BTG-v/- (G8) x B73 WT cross produced progeny with reduced colour expression, whereas progeny from the BTG-a/- x BTG-v/- (G6) cross produced a proportion of plants exhibiting variegated and full colour expression (the numbers of which were not recorded) (Fig.66).

Collectively, these data indicate that the variegated expression of BTG is stable through meiosis.

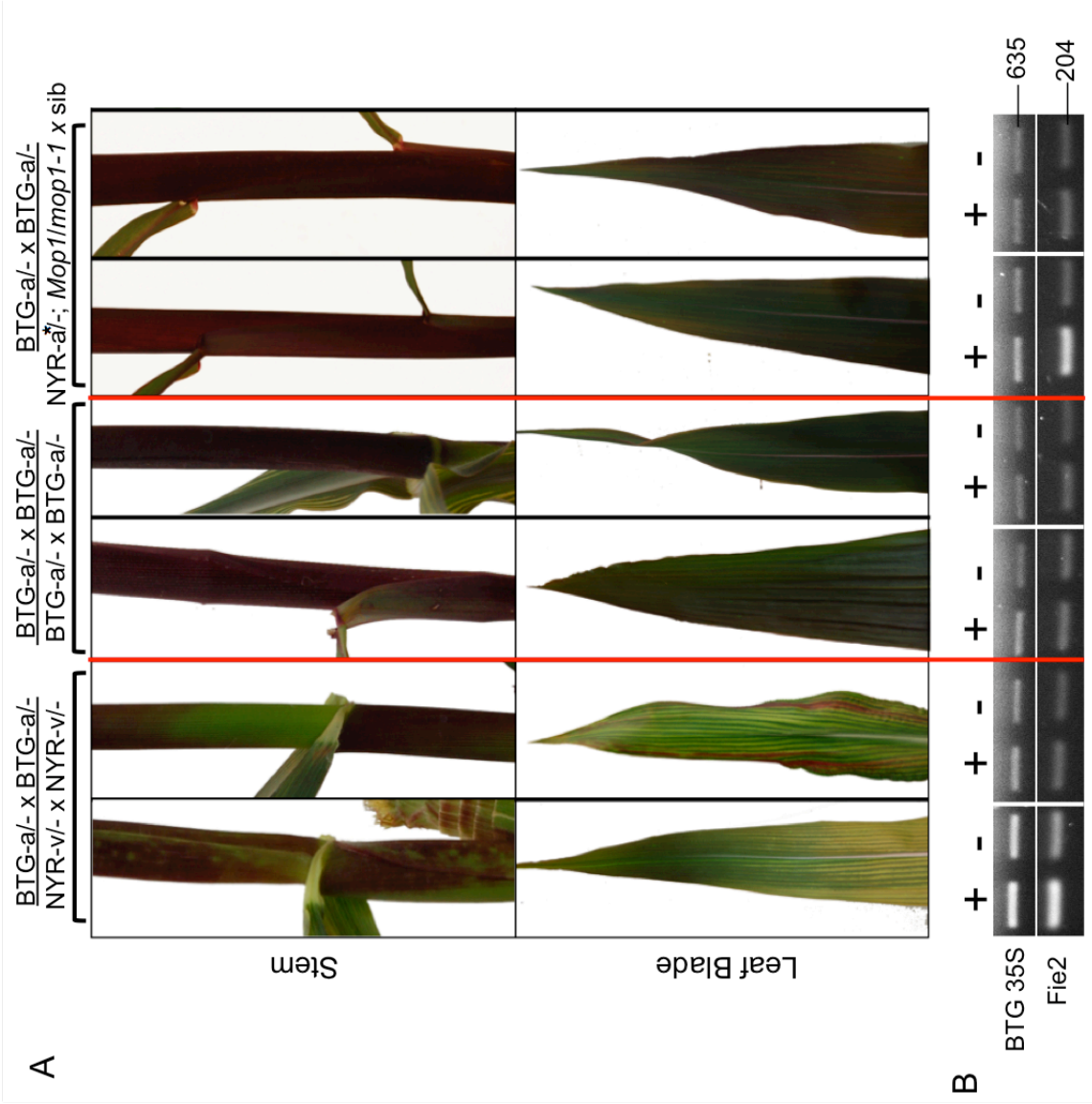


Figure 64: Phenotypic and DNA methylation analysis of grafted plants. (A) Stem and leaf 4 blade pigmentation of two plants displayed variegated pigmentation derived from growth with NYR-v endosperm, two BTG-a grafted plants grown with BTG-a endosperm and two BTG-a grafted plants grown with NYR-v *mop1-1*. Graft labeled as embryo genotype / endosperm genotype. (B) McrBC digestion and PCR amplification of BTG 35S, NYR e35S promoters and *Fie2* (unmethylated control). PCR product sizes (in bp) is indicated on right hand side of the panel.

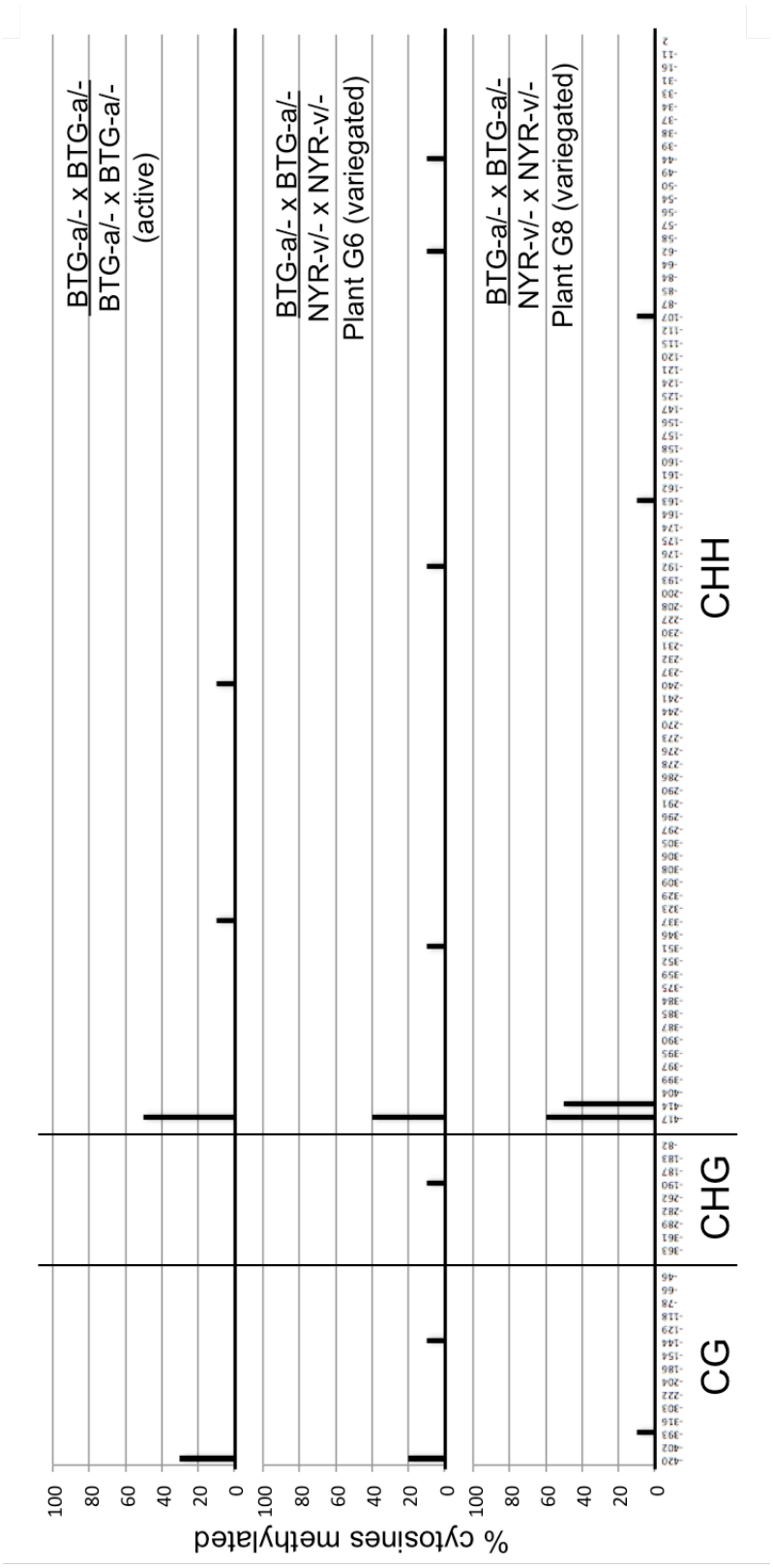


Figure 65: Methylation analysis of BTG 35S promoter of plants derived from endosperm and embryo grafting. Bisulfite analysis of DNA methylation in the BTG promoter of grafted BTG-a control plant and two grafted plants displaying variegated pigmentation. The level of methylation is reported as the percentage of individual cytosine locations exhibiting methylation, established from 10 independent clones, spanning -420 to +2 of BTG (cytosine positions marked on the x-axis and have been split between CG, CHG and CHH contexts), a total of 107 cytosines. Lollipop diagram of methylation in Appendix Fig.77.

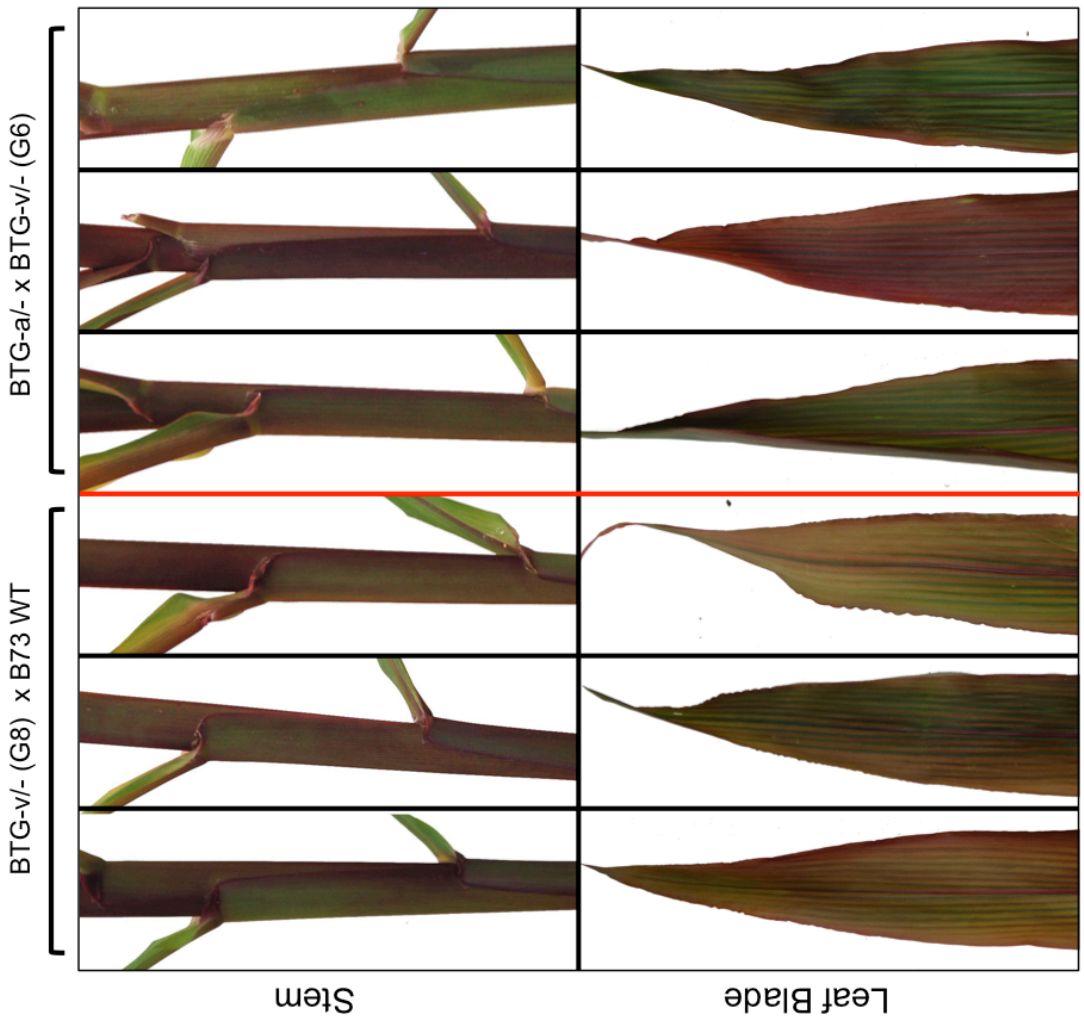


Figure 66: Phenotypic analysis of pigmentation in the F1 progeny from endosperm/embryo grafted plants. Stem and Leaf pigmentation of three plants derived from two independent cross events using grafted plants showing variegated expression G6 and G8. Representative of the range of plants grown.

5.3 Discussion

5.3.1 *Trans*-silencing of BTG-a by NYR-v

My results so far have shown that NYR-v epigenetic silencing is transmittable to BTG-a upon genetic introgression, confirming one of the aims of the chapter. This initially produced a variegated expression of BTG in progeny plants, reminiscent of the expression patterns observed of NYR-v by confocal microscopy. This BTG phenotype has previously been observed in a study of epigenetic regulation of transgene silencing utilising this transgene (BTG) and was in part attributed to methylation of the 35S promoter (Madzima et al. 2011).

The sequence homology between the NYR-v and BTG-a promoters is thought to initiate *trans*-silencing, as observed in many other transgene systems (Park et al. 1996; Meyer et al. 1993; Khaitová et al. 2011). Furthermore, the identification of 24 nt small RNAs and DNA methylation located at the NYR-v promoter suggests that methylation can be directed to the BTG-a promoter. Past examples of *trans*-silencing have often shown the transferal of epigenetic marks, such as DNA methylation from one region of homology to another (Meyer et al. 1993; Khaitová et al. 2011). However, studies involving sequence homology based silencing of 35S promoter, have often also concluded that PTGS is the cause, as a result of the high level of transcription by the 'RNA threshold model' (Fagard and Vaucheret 2000; Lindbo et al. 1993). But, as previously stated homology only exists between the promoters of the two transgenes and not their coding sequences, therefore RdDM TGS is more likely.

DNA methylation analysis of two BTG individuals presenting strong and weak silencing following NYR-v introgression showed very different levels. The NYR-v/-; BTG-v/- (strong silencing) individual possessed very high levels of methylation, including all cytosines within the *asf-1* element and much of the CAAT-like elements. On the other hand, the NYR-v/-; BTG-v/- (weak silencing) individual showed next to no methylation at all, and no detectable methylation within any of the functional 35S positions previously mentioned. The latter profile was also indistinguishable from the lack of methylation observed at the promoter of the BTG-a control plant. In addition, no mixed methylation clones were observed in the bisulfite sequencing data that may account for variegated expression in the BTG-v (strong silencing) individual, and the same is true of the BTG-v (weak silencing) individual. However, this may only be a result of the analysis only sequencing 10 clones, which may not have been enough to find promoter profiles of differential methylation, as both the variegated plants analysed showed extremes of BTG expression. It would be advantageous to have carried out methylation analysis upon a BTG-v plant showing more moderate levels of expression, to observe if this occupied a middle ground between the two current data sets. McrPCR analysis

of a BTG-v plant with intermediate expression did show low methylation compared to BTG-v (strong silencing). I propose that methylation and silencing of the BTG promoter correlate following introgression with NYR-v, and the evidence of methylation and 24 nt siRNAs identified at NYR-v suggest that BTG-v is brought about by RdDM due to promoter sequence homology.

Interestingly though, the backcrossed (BC) generation of BTG plants generated, with and without NYR-v, produced completely silent BTG (BTG-s) plants. This suggests that more efficient silencing had been established at BTG, despite the same silent state not being applied to NYR (still showed variegated expression). Furthermore, the level of DNA methylation and profile of individually sequenced clones at the BTG-s promoter was no different from that seen in the BTG-v (strong silencing) individual of the previous generation. A result similar to this has been observed in the literature in tobacco, investigating paramutation of transgenes (Khaitová et al. 2011). In this study inefficient silencing of a 35S promoter driven reporter was initiated due to introduction of DNA methylation from a silencing element. Over three generations (which segregated out the silencing element) the level of methylation remains relatively constant, however repression intensifies, resulting in full silencing of the reporter (Khaitová et al. 2011). A similar effect could be responsible for the reinforcement of silencing producing BTG-s expression levels, despite loss of NYR-v. This is not a new observation, as silencing of transgenes over multiple generations has been shown in the past, often associated with increasing levels of DNA methylation (Meyer and Saedler 1996; Assaad et al. 1993; Kilby et al. 1992). However, as previously stated no difference in methylation was detected between BTG-s and BTG-v (strong silencing) plants. This may be the result of the number of clones sequenced for analysis of BTG-v (strong silencing), as mentioned before, or this may be the result of other epigenetic changes, such as the efficient attraction of repressive histone modifications by DNA methylation. Comparison of individual clones sequenced between BTG-s and BTG-v (strong silencing) individuals also does not yield much information, with only a small reduction detected at a single cytosine position within the *asf-1* region in the BTG-v individual. However, NYR-v/-; BTG-v/- (weak silencing) plants outcrossed also became BTG-s, and initial analysis showed this BTG-v variation to possess little DNA methylation. It remains unclear how this transition has occurred, despite the low levels of silencing observed in BTG-v (weak silencing) individuals (Fig.50).

The emergence of BTG-s following segregation of NYR-v and further crossing of -/-; BTG-s/- into CYR-a plants showed that BTG-s state is epigenetically stable. The methylation profile for this last generation, however, was not investigated. Also the loss of NYR-v has not changed the sequence contexts of methylation found at the BTG-s promoter, a high abundance at CHH positions are

still methylated, indicating the action of *de novo* methylation by RdDM (Law and Jacobsen 2010). In order for CHH methylation to be propagated in the next generation it would require active signaling, and cannot be guided by hemimethylated daughter strands (Law and Jacobsen 2010). Therefore, small RNAs required for signaling may have been 'unlocked' at the BTG-s promoter, allowing for continued silencing, and therefore also suggest that paramutagenic abilities may be possible.

5.3.2 BTG-s shows paramutation-like ability

Crosses between BTG-s (previously silenced by NYR-v) were able to initiate silencing of the CYR-a transgene, generating offspring which showed variegated expression (CYR-v) (Fig.67). But the number of plants exhibiting this expression profile (BTG-s/-; CYR-v/-) was only 7 plants out of potentially 20 total with this genotype in the experiment performed. My initial investigation germinating a smaller number of seeds failed to find these events, indicating that the silencing is occuring at low frequencies. The reason for this may be the result of the small difference in promoter sequence between BTG and CYR resulting in a stochastic action of silencing, however, this seems unlikely due to the ability of NYR-v (same e35S promoter as CYR) to silence BTG-a. Alternatively, the result may be due to the lack of tandem element in the BTG transgene, leading to the production of fewer small RNAs that can facilitate RdDM at CYR-a. However, this hypothesis assumes that BTG-s is generating site specific 24 nt siRNAs and also that methylation has been directed to CYR, both of which have not been shown. The current model for paramutation, as proposed by Arteaga-Vazquez and Chandler (2010) uses *b1* paramutation as a working base, stating the essential requirement of multiple tandem repeat elements and components of the RdDM pathway. This model indicates two possible roles for tandem repeats, functioning as the source of small RNAs to pass silencing signal in *trans* and also the possibility of DNA-DNA interactions between distant loci (Alleman et al. 2006; Arteaga-Vazquez and Chandler 2010). The lack of tandem repeats in BTG, however, suggest that paramutation-like events do not always require this genetic structuring, and the sustenance of CHH methylation at the BTG-s promoter (as previously stated) suggests siRNAs are functional at this location. Additionally, the action of RdDM and 24 nt siRNAs by NYR-v is reinforced by the failure to silence CYR-a in an RdDM mutant background (data not shown).

The only other examples of transgene paramutation in plants have shown *trans*-silencing and not true paramutation, which involves the production of a paramutagenic epiallele following segregation of the original silencing allele (Khaitová et al. 2011). Transgene paramutation has been

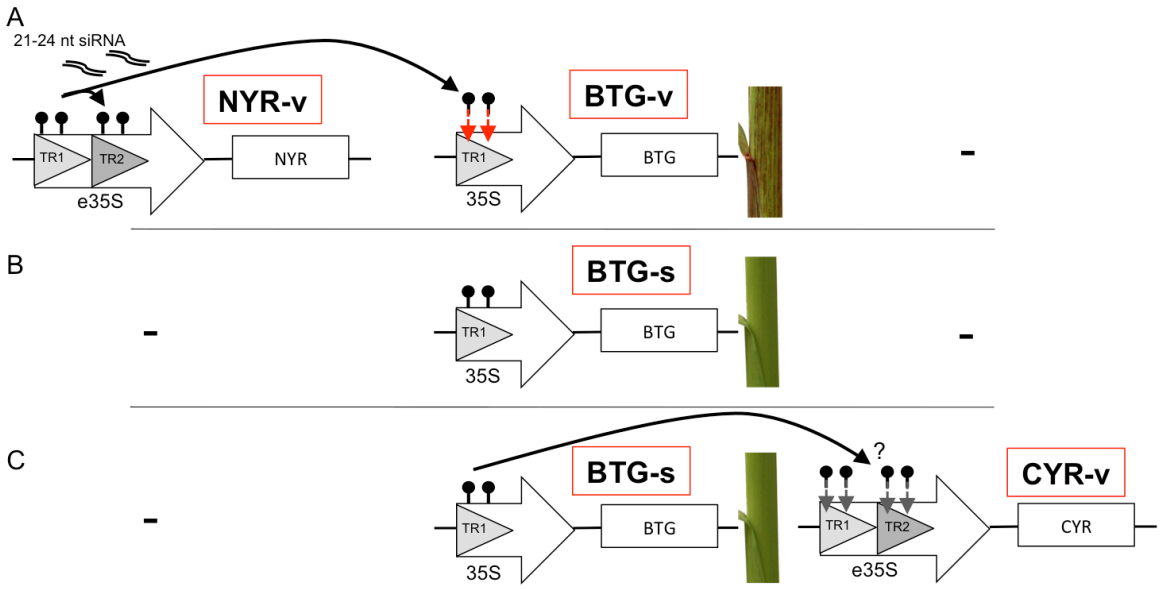


Figure 67: Schematic diagram of transgene expression states to show paramutation. Three consecutive generations (A-C), showing (A) introgression of NYR-v and BTG-a (NYR-v/-; BTG-v/-), (B) outcrossing of BTG-v to separate from NYR-v (-/-; BTG-s/-) and (C) introgression of BTG-s with CYR-a (CYR-a/-; BTG-s/-). Black lollipops represent methylation, red arrows represent action of *de novo* methylation, black arrows represent activity of 24 nt siRNAs, grey arrows represent potential direction of methylation and stem images show BTG phenotype in the plant.

achieved, however, by use of transgenes with endogenous factors known to be paramutagenic, such as the *P1-rr* allele which encodes transcription factors for phlobaphene biogenesis that produces a paramutagenic state in endogenous *P1-rr* (Sidorenko et al. 2001). The data presented in this report have shown that NYR-v can transmit silencing in *trans* to BTG-a, which, can in turn silence CYR-a in a paramutagenic-like manner. This is despite BTG lacking a tandem repeat at its promoter, the sequence of homology between all three transgenes. BTG-s appears to retain its postulated small RNAs involved in RdDM, as evidenced by the reinforcement of asymmetric methylation without tandem repeats.

5.3.3 Variegation of BTG expression in grafted plants

Another primary aim for this chapter was to utilise NYR-v to investigate small RNA movement in the seed. The hypothesis of small RNA movement between the endosperm and embryo of seeds has been derived from the observation of genome demethylation in the endosperm (Hsieh et al. 2009). This suggests that the endosperm can act as a sacrificial tissue in directing embryo silencing through an epigenetic contribution of siRNAs (Mosher and Melnyk 2010). Unfortunately, utilising heterofertilisation to assess small RNA movement in the seed failed to produce significant data.

My experiment may not have been large enough in order to detect a single heterofertilisation event, occurring at a frequency of between 0.5-5%. However, the frequency of events are also known to vary, down to a frequency as little as 0.14%, (Gao et al. 2011) detected in experiments that used a much larger population size than my investigations (see 5.1.5). A recent study in maize may aid in the ability to detect these events, showing that heterofertilised seeds produce smaller embryos, providing an easier method for the identification of these rare events and also giving biological insight into the importance of genetically concordant seed tissues (Wu et al. 2013). Furthermore, it has been recently shown in *Arabidopsis* that disrupting the *fertilisation-independent seed* (*fis*) pathway leads to heterofertilisation events (Maruyama et al. 2013), thus similar experiments could be carried out in *Arabidopsis* in the future.

Moreover, the levels of NYR small RNAs were not measured and due to the nature of the endosperm may behave differently to leaves. This is especially relevant due to the different NYR endosperm phenotypes observed in 3.2.9, due to differential parental transmission orientations. Work carried out showing movement of small RNAs between root and shoot tissue in *Arabidopsis* used a transgene to generate large quantities, far more than are probably derived from endogenous regions (Melnyk et al. 2011), and far more than the 58.14 RPM measured from NYR-v in vegetative tissue. Finally, if movement between endosperm and embryo was achieved in this experiment it is also unknown whether an epigenetic change of BTG resulting in silencing would be inherited mitotically by the developing embryo. It is possible that an epigenetic change may not be somatically stable and thus undetectable in my screen.

The 2nd experiment exploring small RNA movement used tissue grafting of genetically dissimilar endosperm and embryo and generation of a mature plant. This technique has not been utilised before, but my experiments have shown that it is sufficient to develop mature plants from immature embryos placed upon endosperm tissue. It provides a novel way to study the interaction between endosperm and embryo. Out of 41 plants, two displayed variegated pigmentation. The two individuals exhibiting this phenotype were both derived from BTG-a embryos grafted to NYR-v endosperm and showed very low levels of methylation at the promoter, which did not differ from the BTG-a control in abundance, only location. Furthermore, the variegated BTG expression state is heritable as a subsequent generations of individuals have shown. An experiment was also planned here to investigate if grafted BTG-v crossed with BTG-a would result in silencing, but this has not been followed up. However, methylation analysis of NYR-v/-; BTG-v/- (weakly silenced) also showed methylation levels equal to this, and therefore it may be that the 10 clones sequenced was sufficient to show a mixed population of profiles. It is also possible that methylation may

only provide part of the answer to the epigenetic marks involved in controlling silencing of the transgenes. The lack of extensive methylation in the variegated grafted individuals may indicate that heritable chromatin alterations may be associated with silencing.

This technique may be suitable to investigate the regulation of endogenous genes in the same manner as explored using transgenes, as demonstrating this phenomenon between transgenes is very different to showing its occurrence between endogenous elements. The introduction of a transgene with sequence homology to an endogenous gene, to suppress its expression, within endosperm tissue could be used in an attempt to silence the genes expression in the embryo/mature plant.

5.3.4 Silencing is not co-localised between NYR and CYR

A preliminary analysis of the spatial expression of NYR-v and CYR-v within root tissue established that active expression of the two transgenes do not always overlap. This result is unexpected, as the action of silencing within a single cell, in the case of a *trans*-silencing event, would logically be thought to act equally on both transgenes. Therefore, both transgenes should be silent/active within a single cell, but instead expression is stochastic. The differences in silencing patterns may be attributed to differential chromatin states of the transgenes insertion locations or it may be an effect of differences between *cis* (of NYR-v) and *trans* (of CYR-v) silencing. Less NYR expressing individual cells are observed compared with CYR, suggesting that *cis* silencing of NYR is stronger. These data raise questions about the modes of initiation and maintenance of silencing of different loci within the cell, as much about variegated expression is still not understood (Brabbs et al. 2013; Marenkova et al. 2012). This is discussed in greater detail in section 6.1.

5.3.5 Summary

NYR-v was found to silence the BTG-a transgene in *trans*, directing *de novo* DNA methylation changes at the 35S promoter, likely to be responsible for variegated BTG pigmentation. In the absence of NYR-v, full silencing of BTG (BTG-s) was adopted, associated with DNA methylation at the promoter.

BTG-s segregated from NYR-v, was also found to silence CYR-a in a paramutation-like action producing variegated CYR individuals.

The grafting of NYR-v derived endosperm and BTG-a embryos generated variegated BTG expression in mature plants, which is stable over multiple generations, suggesting the movement of silencing signals between seed components. This technique may be suitable for the epigenetic modification of endogenous genes.

6 General Discussion

6.1 General Discussion

TGS of TEs and repetitive elements allow for the effective co-ordination of the genome, preventing the transcriptional activation of damaging TEs and retaining them as a form of epigenetic memory against future exogenous insertion (Lisch 2009). However, the nature of complex plant genomes, such as maize presents a much greater abundance of TEs and repetitive elements than plants with simpler genomes (Schnable et al. 2009). This genome conformation must therefore possess different functioning epigenetic mechanisms acting at TEs and is also likely to effect the expression of genes, due to the close proximity of the greater abundance of TEs (Eichten et al. 2012). The ability of the genome to regulate expression is still not well understood, especially outside of the model plant species *Arabidopsis*. Currently the RdDM pathway is established is the means to repress non-coding regions of the genome, using siRNAs to direct the placement of DNA methylation, resulting in the installation of a heterochromatic state, initiating gene silencing. The reasons for the difference in genome architecture is still under discussion, but it has been proposed that complex genomes may allow for rapid genome restructuring allowing for greater evolutionary flexibility (Fedoroff 2012).

The work in this thesis has analysed the mechanisms involved in the silencing of the NYR-v transgene to investigate gene silencing in maize. Interestingly, this analysis has revealed that despite 24 nt siRNAs and DNA methylation present at the NYR-v promoter and reactivation in RdDM mutant backgrounds, no clear correlation between the presence of methylation and expression could be drawn in vegetative or reproductive tissues. Furthermore, mutants of RdDM components resulted in only a proportion of progeny being reactivated (NYR-a*). However, previous studies have shown that DNA methylation and expression can be independent of each other (Amedeo et al. 2000; Dieguez et al. 1998; Scheid et al. 1998). This is also evident following restoration of *rmr1* following outcrossing with WT where plants still showed hypomethylated promoter regions, reminiscent of the mutant background, but maintained a NYR-a* state.

Furthermore, the proportion of offspring resulting in reactivation and the presence of methylation despite loss of RdDM function also suggests that other mechanisms other than RdDM are involved in silencing. This is also proposed by the reactivation utilising *rmr1-1*, suspected to have association with chromatin marks. Analysis by Chromatin Immunoprecipitation (ChIP) can be used to investigate the presence of histone modifications within specific regions of DNA. Determining if repressive histone modifications are present at NYR-v may provide greater insight into its variegated expression and subsequent reactivation in RdDM mutant backgrounds. A study by Amedeo et al. (2000) showing methylation independent changes in gene expression showed that the presence or absence of H3K9me2 was instead correlated with expression. A similar histone

modification at NYR-v or dynamically effecting its expression from adjacent TEs may confer the silencing observed.

Alternatively, the reactivation of initially a small proportion of NYR-v plants in RdDM mutant backgrounds implicates a partial role for RdDM in silencing. The only site of methylation thought to be responsible for silencing in comparison between NYR-v and NYR-a* plants was methylation at individual positions within the asf-1 region (Kanazawa et al. 2007a). To fully explore the relationship of DNA methylation with NYR-v, mutation of the maintenance methyltransferases, successful chemical removal of methylation or analysis of NYR-a* and NYR-v from the same mutant background would be required. Interestingly, NYR-a* individuals segregated away from *rmr1-1* mutant backgrounds did not revert back to a silenced expression and also showed no correlation between promoter-wide methylation and silencing. Similar maintenance following mutant segregation has been observed previously with *mop1* and *rmr2* mutants (McGinnis et al. 2006). This may represent selection for reactivated epialleles, or the inability for *rmr1* to be involved in establishing silencing. The former raises interesting questions about the coordination of the genome following mutant introgression, if a trait can be progressively selected for reinforcing it, even following reintroduction of the silencing machinery after multiple generations.

The reasons for the establishment of the NYR-v expression are still difficult to explain, however, it is in part dependent on RdDM due to the reactivated progeny generated following their loss of function. The variegated expression in transgenes has been under investigation since the 1990's and much about the dynamic expression traits is still unknown. Therefore expression may be resulted to a probability dependent on a potential myriad of epigenetic marks, which may or may not be dynamic. As previously discussed methylation is still maintained in all sequence contexts at the NYR promoter in RdDM mutant backgrounds. The RdDM-independent DNA methylation placement and potential role of repressive histone modifications therefore contribute as other components of silencing. Furthermore, it remains difficult to determine if a position-effect is the cause of silencing due to the lack of epigenetic analysis of the upstream *gypsy*-like TE and unknown downstream endogenous DNA. The spread of repressive histone modifications and DNA methylation from TEs to flanking coding regions has recently been shown in maize (Eichten et al. 2012). Therefore, comparison of NYR-v with a transgene from a different transformation event, at a different insertion location or the ability to relocate NYR-v to a different position in the genome would be ideal for this investigation. The processes behind the stochastic expression, are as with other transgenes, unclear (Brabbs et al. 2013; Marenkova et al. 2012; Madzima et al. 2011). Position-effects resulting in silencing/variegatied expression are thought to be due to

dynamic changes in the chromatin environment, similar to PEV in *Drosophila* (Tartof et al. 1984). The initiation of silencing may include dynamic chromatin environments allowing in addition to the effects of environmental factors. The later has been studied utilising a *FLC:GUS* fusion in *Arabidopsis*, observing expression levels of *FLC* following different lengths of cold treatments (Angel et al. 2011). Their work revealed that implementation of H3K27me3 at *FLC* (shown by loss of reporter expression) can be quantitatively modeled to determine the proportion of the cell population to have changed epigenetic state, by each treatment. This analysis has observed the control between bistable states, where the switch (vernalisation) is known.

Moreover, observation of individual cells with differing expression states of NYR and CYR was also confusing. This suggests that the stochastic processes of silencing act at each element separately despite the silencing signal originating as a result of the presence of NYR-v. Therefore, the initiation of silencing at CYR-a is not dependent on the expression state of NYR. This may reflect an uneven production of siRNAs originating from NYR-v acting to silence CYR-a, which perhaps do not effect the expression at NYR. This may then also explain the reason for the BTG-v expression upon introduction to NYR-v, followed by BTG-s in subsequent generation where silencing at BTG has been reinforced through the gametes. Without individual analysis of cells of each class it is difficult to draw a firm conclusion on these differences.

Analysis of endosperm and embryo tissue also observed the potential imprinting of NYR-v with parent-of-origin specific expression in the endosperm. Repeat elements and/or proximity to TEs are a hallmark of imprinted genes in plants (Gehring et al. 2009), both attributes possessed by NYR-v. As previously stated if expression observed was due to a gene dosage effect, a uniform level of expression would be observed between all the endosperm observed from each cross orientation, which was not observed. Therefore, it is more likely that different epialleles reliant of methylation at other locations or histone modifications may be the cause as previously discussed.

Intriguingly, 21-22 nt small RNAs mapping to the NYR-v promoter were also detected. This class of small RNAs have recently been shown to direct de novo DNA methylation by a novel Pol II - RDR6 pathway, as well as acting in PTGS (Nuthikattu et al. 2013). It has been traditionally thought that TGS and PTGS were independently operating mechanisms, but this may not be the case. The traditionally roles of TGS and PTGS are also being challenged, with new work showing that PTGS can induce meiotically heritable changes in gene expression by influencing siRNA biogenesis (Zhong et al. 2013). Also PTGS has also been shown to lead to TGS over multiple generations (Khaitová et al. 2011). This suggests that gene silencing relies on both of these systems, perhaps in sequence, to induce transcriptional gene silencing.

Interestingly, following *trans*-silencing of BTG-a by NYR-v, methylation at the promoter of BTG-v showed correlation with silencing. This suggests that different forms of silencing have occurred at each of the transgenes, perhaps dependent on the insertion location, specifically of NYR-v. The initial attraction of NYR-v silencing may be locus/sequence specific, hence the silencing of a single NYR transformation event, and not others generated. The behavior of NYR-v may be indicative of endogenous genes in maize, and has shown the possibility of regulation greater than a single pathway that we are currently aware. Just as many other studies have shown transgenes behaving in independent ways has been concluded to be locus/sequence specific effects (McGinnis et al. 2006; Madzima et al. 2011). Whether this transgene can behave in the same manner in plants with simpler genomes is unknown, as no previously studied transgene I am aware of has shown this combination of traits.

It is also important to consider the genomes reaction to the insertion of foreign DNA such as transgenes. Observation may therefore display effects towards DNA of similar structure, and not endogenous genes. Work conducted in tobacco utilised a transgene designed to mimick the structure of endogenous genes, enlisting an endogenous promoter, 5' and 3' UTRs, introns and terminator elements while encoding expression for GFP (Dadami et al. 2013). Following transformation, silencing was found to establish itself more quickly within a conventional transgenes than the transgene mimicking an endogenous gene, despite equal levels of methylation detected between both (Dadami et al. 2013). Therefore the regulation of NYR-v may not be typical of endogenous genes due to a differential structure, further identifying them as foreign elements. However, the behavior of NYR-v observed may provide insight into the genomes response to foreign elements, potentially of viral origin. This is especially relevant for plants with complex genomes in understanding the mechanisms utilised in accumulating TEs and regions of non-coding repetitive DNA (Fedoroff 2012).

Genome complexity may also contribute to the abundance of paramutation events, whereby the epigenetic landscape of TEs may provide transcriptional control over genes. Repetitive elements and TEs have been implicated in the cause of paramutation, since the discovery of an upstream seven tandem repeat sequence was found to be required for *b1* paramutation (Hollick 2012). Furthermore, TEs have also been implicated in maintaining the silenced expression of *pl1* following increased transcription when introduced into a Pol IV (*rmr6-1*) mutant background (Erhard et al. 2009). However, conserved mechanisms enabling paramutation have eluded investigation, producing multiple theories. The most prominent of which, as stated in the introduction (see 1.2.3), is that RdDM can facilitate changes in DNA methylation between regions of homology on dif-

fering chromosomes dependent on targeting by siRNAs (Mette et al. 2000). On the other hand, an alternate hypothesis describes the 'pairing' model where physical interaction between the two epialleles can occur, similar to observations in *Drosophila* (Henikoff 1997). However, a growing body of evidence has called the role role of siRNAs into question, due to the surprising findings that *rmr1-1* and *rmr2-1*, which both decrease siRNA abundance are not required for paramutation in some instances Hale et al. 2007; Barbour et al. 2012). This is in contrast to other components of the RdDM pathway, such as *mop1* that are essential (Dorweiler et al. 2000). The requirement of repeat DNA allowing paramutation is also confusing, thought to provide a source of small RNAs to act in trans, however, small RNAs detected from the tandem repeats of B' and B-I epialleles do not differ, despite differing methylation levels (Hollick 2012). Additionally, analysis of this region also showed that H3K9 and H3K27 methylation are also associated with the the heavily methylated epiallele (Haring et al. 2010).

Currently paramutation has not been shown in transgenes, unless they possess previously known paramutagenic endogenous genes. Recent investigations into paramutation-like *trans*-silencing of SALK T-DNA has revealed silencing between T-DNA insertions and an endogenous gene, oddly resulting in increased expression (Xue et al. 2012). However, the paramutagenic T-DNAs must be homozygous following reversion of expression when outcrossed to WT. Therefore, this paramutation-like event is difficult to distinguish from complicated trans interactions as components have not been separated from each other following change in expression. This study also does not utilise any repeats, thought to be required for paramutation. The paramutation-like effects observed following NYR-v *trans*-silencing of BTG-a and subsequently CYR-a also disagree with the requirement for repeats. Both NYR-v and CYR-a possess tandem repeats within the promoter region, whereas BTG-a does not. But, BTG-s was able to silence expression of CYR-a, albeit at a much lower proportion of plants than observed between the first silencing event (NYR-v and BTG-a). Moreover, all three transgenes are different, located at distinct locations in the genome with different neighbouring endogenous sequences. Hollick (2012) suggests that nearby TEs familiar to differing epialleles in maintaining a transcriptional silenced state. This differs from the observation of paramutation of endogenous genes where surrounding DNA is largely the same, unless changed due to recombination events. The differing genomic landscapes present proposes that it is only interaction between the homology of the transgene sequences, such as the homology of the promoters that results in the paramutation-like results observed.

Although the majority of research in paramutation has occurred in maize recent work crossing two *Arabidopsis* ecotypes Lansberg *erecta* and C24 has revealed trans-chromosomal methylation

that resembles paramutation-like effects (Greaves et al. 2012). The authors hypothesised that this mechanism contributes to the vigor found in these *Arabidopsis* hybrids, however, the vigor diminished over subsequent generations. This preliminary analysis still requires greater investigation to determine if the differences in small RNA and DNA methylation observed is attributable to paramutation.

Paramutation events are still quite rare, only a few endogenous examples are known, however, the potential reliance on siRNAs, known to be mobile (Melnyk et al. 2011) has suggested that movement might also occur intercellularly.

The intercellular movement of silencing signals also led to analysis of potential intercellular epigenetic communication between developing endosperm and embryo tissue. Intercellular communication by small RNAs has been shown between vegetative tissues and within pollen, and the demethylated landscape of the endosperm suggests its existence between endosperm and embryo tissues (Mosher and Melnyk 2010). My initial experiment, taking advantage of the heterofertilisation phenomenon in maize, was unsuccessful in attempting to demonstrate this. However, I have shown that a silencing phenotype can be passed from grafted endosperm with NYR-v to developing embryos with BTG-a, generating BTG-v individuals, thought to be dependent on siRNAs. Furthermore, methylation analysis of variegated grafted plants showed a very low levels present, similar to the profile observed in the NYR-v/-; BTG-v/- (weak silencing) plant. This suggests that other epigenetic changes, as yet unknown, must be responsible, which may include repressive histone modifications or direction of DNA methylation at other locations. The latter seems unlikely due to the only sequence of homology existing between the transgenes residing at the promoter, and as stated previously methylation may be independent of silencing. As stated previously though methylation and expression can be independent of each other. Furthermore the maintenance of a variegated expression to the subsequent generation of grafted plants shows that silencing is meiotically heritable. But, this has not behaved in the same manner as -/-; BTG-s/- plants which resulted in fully silenced BTG plants, not propagation of the variegated expression. This may be a result of the inheritance of a weaker silencing signal or the lack of NYR-v for a whole generation with BTG-a (NYR-v/-; BTG-v/- plants).

The recent discoveries of small RNA movement between other reproductive components, such as the vegetative nuclei and sperm cells (Slotkin et al. 2009) and the central cell and egg cell (Ibarra et al. 2012), indicates a communication network reliant on movement of epigenetic information in the form of small RNAs derived from companion cells. The implications of epigenetic communication between seed components are highly significant, as environmental factors could lead to the

expression of small RNAs that could contribute to novel transgenerational memory transmitted to the offspring. The findings of this report have demonstrated that communication of silencing between seed components is possible.

6.2 Conclusions

The analysis of NYR-v has revealed that gene silencing in maize can involve more than a single established pathway, such as RdDM. Although the exact modifications present at NYR-v required for silencing were not determined, loss of the RdDM pathway did result in reactivation in addition to *rmr1-1* mutants not available in Arabidopsis.

Utilising NYR-v I was able to successfully demonstrate paramutation-like events through silencing of the BTG-a and CYR-a transgenes and the ability to make them paramutagenic. This system was then also used for grafting genetically dissimilar endosperm and embryo tissue, presenting a barrier to NYR-v derived silencing of BTG-a that was overcome with silencing signal transmitted to/sensed by the mature plant. This reveals requirement currently used to define paramutation of endogenous genes may not be essential and that small RNA movement between seed components is very likely.

6.3 Future Perspectives

The work conducted here has focused on the use of the NYR-v transgene, however, as described this has limitations and the study of endogenous elements require investigation to determine the processes involved in their regulation in the genome. Furthermore, the differences in mechanisms involved in generating complex genomes may be provided as well as this.

Furthermore, the paramutation-like effects contributed by NYR-v require further investigation to determine the means by which silencing was transmitted to BTG-a and CYR-a transgenes in order to better understand how endogenous paramutation might function.

Finally, the epigenetic signaling of silencing between seed components needs to be explored, determining if small RNAs are specifically translocated and if evolutionary advantages can be conveyed to the zygote and mature plant from the sacrificial endosperm tissue during seed development.

References

Alieva, N. O., K. A. Konzen, S. F. Field, E. A. Meleshkevitch, M. E. Hunt, V. Beltran-Ramirez, D. J. Miller, J. Wiedenmann, A. Salih, and M. V. Matz (2008) 'Diversity and evolution of coral fluorescent proteins.' *PLoS One* 3(7), e2680

Alleman, M., L. Sidorenko, K. McGinnis, V. Seshadri, J. E. Dorweiler, J. White, K. Sikkink, and V. L. Chandler (2006) 'An RNA-dependent RNA polymerase is required for paramutation in maize.' *Nature* 442(7100), 295–298

Allen, E., Z. Xie, A. M. Gustafson, and J. C. Carrington (2005) 'microRNA-directed phasing during trans-acting siRNA biogenesis in plants.' *Cell* 121(2), 207

Amedeo, P., Y. Habu, K. Afsar, O. Mittelsten Scheid, and J. Paszkowski (2000) 'Disruption of the plant gene MOM releases transcriptional silencing of methylated genes.' *Nature* 405(6783), 203–206

Angel, A., J. Song, C. Dean, and M. Howard (2011) 'A Polycomb-based switch underlying quantitative epigenetic memory.' *Nature* 476(7358), 105–108

Armstrong, C. L. (1994) 'Regeneration of plants from somatic cell cultures: applications for the in vitro genetic manipulation', vol. 663-671 (The Maize Handbook, New York)

Arteaga-Vazquez, M. A., and V. L. Chandler (2010) 'Paramutation in maize: RNA mediated trans-generational gene silencing.' *Current Opinion in Genetics and Development* 20(2), 156–163

Arteaga-Vazquez, M. A., L. Sidorenko, F. A. Rabanal, R. Shrivistava, K. Nobuta, P. J. Green, B. C. Meyers, and V. L. Chandler (2010) 'RNA-mediated trans-communication can establish paramutation at the b1 locus in maize.' *PNAS* 107(29), 12986–12991

Ashe, A., A. Sapetschnig, E. Weick, J. Mitchell, M. P. Bagijn, A. C. Cording, A. Doebley, L. D. Goldstein, N. J. Lehrbach, and J. Le Pen (2012) 'piRNAs can trigger a multigenerational epigenetic memory in the germline of C. elegans.' *Cell* 150(1), 88–99

Assaad, F. F., K. L. Tucker, and E. R. Signer (1993) 'Epigenetic repeat-induced gene silencing (RIGS) in Arabidopsis.' *Plant Molecular Biology* 22(6), 1067–1085

Ausin, I., T. C. Mockler, J. Chory, and S. E. Jacobsen (2009) 'IDN1 and IDN2 are required for de novo DNA methylation in Arabidopsis thaliana.' *Nature Structural and Molecular Biology* 16(12), 1325–1327

Autran, D., C. Baroux, M. T. Raissig, T. Lenormand, M. Wittig, S. Grob, A. Steimer, M. Barann, U. C. Klostermeier, and O. Leblanc (2011) 'Maternal epigenetic pathways control parental contributions to Arabidopsis early embryogenesis.' *Cell* 145(5), 707–719

Bannister, A. J., P. Zegerman, J. F. Partridge, E. A. Miska, J. O. Thomas, R. C. Allshire, and T. Kouzarides (2001) 'Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain.' *Nature* 410(6824), 120–124

Barbour, J. R., I. T. Liao, J. L. Stonaker, J. P. Lim, C. C. Lee, S. E. Parkinson, J. Kermicle, S. A. Simon, B. C. Meyers, and R. Williams-Carrier (2012) 'required to maintain repression2 is a novel protein that facilitates locus-specific paramutation in maize.' *Plant Cell* 24(5), 1761–1775

Baubec, T., A. Pecinka, W. Rozhon, and O. M. Scheid (2009) 'Effective, homogeneous and transient interference with cytosine methylation in plant genomic DNA by zebularine.' *The Plant Journal* 57(3), 542–554

Baubec, T., H. Q. Dinh, A. Pecinka, B. Rakic, W. Rozhon, B. Wohlrab, A. von Haeseler, and O. M. Scheid (2010) 'Cooperation of multiple chromatin modifications can generate unanticipated stability of epigenetic states in Arabidopsis.' *Plant Cell* 22(1), 34–47

Baucom, R. S., J. C. Estill, C. Chaparro, N. Upshaw, A. Jogi, JM. Deragon, R. P. Westerman, P. J. SanMiguel, and J. L. Bennetzen (2009) 'Exceptional diversity, non-random distribution, and rapid evolution of retroelements in the B73 maize genome.' *PLoS Genetics* 5(11), e1000732

Becraft, P. W., and J. F. Gutierrez-Marcos (2012) 'Endosperm development: dynamic processes and cellular innovations underlying sibling altruism.' *Wiley Interdisciplinary Reviews: Developmental Biology* 1(4), 579–593

Benfey, P. N., and N. Chua (1990) 'The cauliflower mosaic virus 35S promoter: combinatorial regulation of transcription in plants.' *Science* 250(4983), 959–966

Berger, S. L. (2007) 'The complex language of chromatin regulation during transcription.' *Nature* 447(7143), 407–412

Bièche, I., M. Olivi, M. H. Champème, D. Vidaud, R. Lidereau, and M. Vidaud (1998) 'Novel approach to quantitative polymerase chain reaction using real-time detection: application to the detection of gene amplification in breast cancer.' *International Journal of Cancer* 78(5), 661–666

Bivalkar-Mehla, S., J. Vakharia, R. Mehla, M. Abreha, J. R. Kanwar, A. Tikoo, and A. Chauhan (2011) 'Viral RNA silencing suppressors (RSS): novel strategy of viruses to ablate the host RNA interference (RNAi) defense system.' *Virus research* 155(1), 1–9

Borsani, O., J. Zhu, P. E. Verslues, R. Sunkar, and J. Zhu (2005) 'Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis.' *Cell* 123(7), 1279–1291

Bostick, M., J. K. Kim, PO. Estève, A. Clark, S. Pradhan, and S. E. Jacobsen (2007) 'UHRF1 plays a role in maintaining DNA methylation in mammalian cells.' *Science* 317(5845), 1760–1764

Brabbs, T. R., Z. He, K. Hogg, A. Kamenski, Y. Li, K. H. Paszkiewicz, K. A. Moore, P. O'Toole, I. A. Graham, and L. Jones (2013) 'The stochastic silencing phenotype of Arabidopsis morc6 mutants reveals a role in efficient RNA-directed DNA methylation.' *The Plant Journal* 75, 836–846

Brenner, S., M. Johnson, J. Bridgham, G. Golda, D. H. Lloyd, D. Johnson, S. Luo, S. McCurdy, M. Foy, M. Ewan, R. Roth, D. George, S. Eletr, G. Albrecht, E. Vermaas, S. R. Williams, K. Moon, T. Burcham, M. Pallas, R. B. DuBridge, J. Kirchner, K. Fearon, J. Mao, and K. Corcoran (2000) 'Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays.' *Nature Biotechnology* 18(6), 630–634

Brink, R. A. (1956) 'A genetic change associated with the R locus in maize which is directed and potentially reversible.' *Genetics* 41(6), 872

Brosnan, C.A., and O. Voinnet (2011) 'Cell-to-cell and long-distance siRNA movement in plants: mechanisms and biological implications.' *Current Opinion in Plant Biology* 14(5), 580–587

Bubner, B., and I. T. Baldwin (2004) 'Use of real-time PCR for determining copy number and zygosity in transgenic plants.' *Plant Cell Reports* 23(5), 263–271

Buckley, B. A., K. B. Burkhart, S. G. Gu, G. Spracklin, A. Kershner, H. Fritz, J. Kimble, A. Fire, and S. Kennedy (2012) 'A nuclear argonaute promotes multigenerational epigenetic inheritance and germline immortality.' *Nature* 489(7416), 447–451

Caldana, C., W. R. Scheible, B. Mueller-Roeber, and S. Ruzicic (2007) 'A quantitative RT-PCR platform for high-throughput expression profiling of 2500 rice transcription factors.' *Plant Methods* 3, 7–7

Cao, J., X. Duan, D. McElroy, and R. Wu (1992) 'Regeneration of herbicide resistant transgenic rice plants following microprojectile-mediated transformation of suspension culture cells.' *Plant Cell Reports* 11(11), 586–591

Cao, X., and S. E. Jacobsen (2002) 'Role of the Arabidopsis DRM methyltransferases in de novo DNA methylation and gene silencing.' *Current Biology* 12(13), 1138–1144

Carlsbecker, A., J. Lee, C. J. Roberts, J. Dettmer, S. Lehesranta, J. Zhou, O. Lindgren, M. A. Moreno-Risueno, A. Vatén, and S. Thitamadee (2010) 'Cell signalling by microRNA165/6 directs gene dose-dependent root cell fate.' *Nature* 465(7296), 316–321

Cedar, H., and Y. Bergman (2009) 'Linking DNA methylation and histone modification: patterns and paradigms.' *Nature Reviews Genetics* 10(5), 295–304

Chaillet, J. R., T. F. Vogt, D. R. Beier, and P. Leder (1991) 'Parental-specific methylation of an imprinted transgene is established during gametogenesis and progressively changes during embryogenesis.' *Cell* 66(1), 77–83

Chalfie, M., Y. Tu, G. Euskirchen, W. W. Ward, and D. C. Prasher (1994) 'Green fluorescent protein as a marker for gene expression.' *Science* 263(5148), 802–805

Chaudhuri, S., and J. Messing (1994) 'Allele-specific parental imprinting of dzr1, a posttranscriptional regulator of zein accumulation.' *PNAS* 91(11), 4867–4871

Chen, M., K. Licon, R. Otsuka, L. Pillus, and T. Ideker (2013) 'Decoupling epigenetic and genetic effects through systematic analysis of gene position.' *Cell Reports* 3(1), 128–137

Chen, X., H. Liang, J. Zhang, K. Zen, and C. Y. Zhang (2012) 'Secreted microRNAs: a new form of intercellular communication.' *Trends in Cell Biology* 22(3), 125–132

Chiang, P. K., R. K. Gordon, J. Tal, G. C. Zeng, B. P. Doctor, K. Pardhasaradhi, and P. P. McCann (1996) 'S-adenosylmethionine and methylation.' *FASEB Journal* 10(4), 471–480

Chiu, W., Y. Niwa, W. Zeng, T. Hirano, H. Kobayashi, and J. Sheen (1996) 'Engineered GFP as a vital reporter in plants.' *Current Biology* 6(3), 325–330

Chopra, S., S. M. Cocciolone, S. Bushman, V. Sangar, M. D. McMullen, and T. Peterson (2003) 'The maize unstable factor for orange1 is a dominant epigenetic modifier of a tissue specifically silent allele of pericarp color1.' *Genetics* 163(3), 1135–1146

Coe Jr, E. H. (1959) 'A regular and continuing conversion-type phenomenon at the B locus in maize.' *PNAS* 45(6), 828

Cokus, S. J., S. Feng, X. Zhang, Z. Chen, B. Merriman, C. D. Haudenschild, S. Pradhan, S. F. Nelson, M. Pellegrini, and S. E. Jacobsen (2008) 'Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning.' *Nature* 452(7184), 215–219

Dadami, E., M. Moser, M. Zwiebel, G. Krczal, M. Wassenegger, and A. Dalakouras (2013) 'An endogene-resembling transgene delays the onset of silencing and limits siRNA accumulation.' *FEBS Letters* 587(6), 706–710

Daxinger, L., T. Kanno, E. Bucher, J. van der Winden, U. Naumann, A. J. M. Matzke, and M. Matzke (2008) 'A stepwise pathway for biogenesis of 24-nt secondary siRNAs and spreading of DNA methylation.' *EMBO Journal* 28(1), 48–57

de Borne Dorlhac, F., M. Vincentz, Y. Chupeau, and H. Vaucheret (1994) 'Co-suppression of nitrate reductase host genes and transgenes in transgenic tobacco plants.' *Molecular and General Genetics* 243(6), 613–621

De Buck, S., P. Windels, M. De Loose, and A. Depicker (2004) 'Single-copy T-DNAs integrated at different positions in the Arabidopsis genome display uniform and comparable beta-glucuronidase accumulation levels.' *Cellular and Molecular Life Sciences* 61(19-20), 2632–2645

De Rubertis, F., D. Kadosh, S. Henchoz, D. Pauli, G. Reuter, K. Struhl, and P. Spierer (1996) 'The histone deacetylase RPD3 counteracts genomic silencing in Drosophila and yeast.' *Nature* 12(384), 589–591

de Wet, J. R., K. V. Wood, D. R. Helinski, and M. DeLuca (1985) 'Cloning of firefly luciferase cDNA and the expression of active luciferase in Escherichia coli.' *PNAS* 82(23), 7870–7873

Dehio, C., and J. Schell (1994) 'Identification of plant genetic loci involved in a posttranscriptional mechanism for meiotically reversible transgene silencing.' *PNAS* 91(12), 5538–5542

Dieguez, M. J., H. Vaucheret, J. Paszkowski, and O. M. Scheid (1998) 'Cytosine methylation at CG and CNG sites is not a prerequisite for the initiation of transcriptional gene silencing in plants, but it is required for its maintenance.' *Molecular and General Genetics* 259(2), 207–215

Dorweiler, J. E., C. C. Carey, K. M. Kubo, J. B. Hollick, J. L. Kermicle, and V. L. Chandler (2000) 'mediator of paramutation1 is required for establishment and maintenance of paramutation at multiple maize loci.' *Plant Cell* 12(11), 2101–2118

Dunoyer, P., C. A. Brosnan, G. Schott, Y. Wang, F. Jay, A. Alioua, C. Himber, and O. Voinnet (2010a) 'An endogenous, systemic RNAi pathway in plants.' *EMBO Journal* 29(10), 1699–1712

Dunoyer, P., C. Himber, and O. Voinnet (2005) 'DICER-LIKE 4 is required for RNA interference and produces the 21-nucleotide small interfering RNA component of the plant cell-to-cell silencing signal.' *Nature Genetics* 37(12), 1356–1360

Dunoyer, P., G. Schott, C. Himber, D. Meyer, A. Takeda, J. C. Carrington, and O. Voinnet (2010b) 'Small RNA duplexes function as mobile silencing signals between plant cells.' *Science* 328(5980), 912–916

Eamens, A., M. Wang, N. A. Smith, and P. M. Waterhouse (2008) 'RNA silencing in plants: yesterday, today, and tomorrow.' *Plant Physiology* 147(2), 456–468

Ebbs, M. L., and J. Bender (2006) 'Locus-specific control of DNA methylation by the Arabidopsis SUVH5 histone methyltransferase.' *The Plant Cell* 18(5), 1166–1176

Ebbs, M. L., L. Bartee, and J. Bender (2005) 'H3 lysine 9 methylation is maintained on a transcribed inverted repeat by combined action of SUVH6 and SUVH4 methyltransferases.' *Molecular and Cellular Biology* 25(23), 10507–10515

Eggleston, W. B., M. Alleman, and J. L. Kermicle (1995) 'Molecular organization and germinal instability of R-stippled maize.' *Genetics* 141(1), 347

Ehrlich, M., M. A. Gama-Sosa, L. H. Huang, R. M. Midgett, K. C. Kuo, R. A. McCune, and C. Gehrke (1982) 'Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells.' *Nucleic Acids Research* 10(8), 2709–2721

Eichten, S. R., N. A. Ellis, I. Makarevitch, C. Yeh, J. I. Gent, L. Guo, K. M. McGinnis, X. Zhang, P. S. Schnable, and M. W. Vaughn (2012) 'Spreading of heterochromatin is limited to specific families of maize retrotransposons.' *PLoS Genetics* 8(12), e1003127

Eichten, S. R., R. Briskine, J. Song, Q. Li, R. Swanson-Wagner, P. J. Hermanson, A. J. Waters, E. Starr, P. T. West, and P. Tiffin (2013) 'Epigenetic and genetic influences on DNA methylation variation in maize populations.' *The Plant Cell* 25(8), 2783–2797

Elgin, S. C. (1996) 'Heterochromatin and gene regulation in Drosophila.' *Current Opinion in Genetics and Development* 6(2), 193–202

Elmayan, T., and H. Vaucheret (1996) 'Expression of single copies of a strongly expressed 35S transgene can be silenced post-transcriptionally.' *The Plant Journal* 9(6), 787–797

Erhard, K. F., J. L. Stonaker, S. E. Parkinson, J. P. Lim, C. J. Hale, and J. B. Hollick (2009) 'RNA polymerase IV functions in paramutation in Zea mays.' *Science* 323(5918), 1201–1205

Eun, C. H., K. Takagi, K. I. Park, M. Maekawa, S. Iida, and K. Tsugane (2012) 'Activation and epigenetic regulation of DNA transposon nDart1 in rice.' *Plant Cell Physiology* 53(5), 857–868

Exner, V., E. Aichinger, H. Shu, T. Wildhaber, P. Alfarano, A. Caflisch, W. Gruissem, C. Köhler, and L. Hennig (2009) 'The chromodomain of LIKE HETEROCHROMATIN PROTEIN 1 is essential for H3K27me3 binding and function during Arabidopsis development.' *PLoS One* 4(4), e5335

Fagard, M., and H. Vaucheret (2000) '(Trans)gene silencing in plants: how many mechanisms?' *Annual Review of Plant Biology* 51(1), 167–194

Fedoroff, N. V. (2012) 'Transposable elements, epigenetics, and genome evolution.' *Science* 338(6108), 758–767

Finer, J. J., P. Vain, M. W. Jones, and M. D. McMullen (1992) 'Development of the particle inflow gun for DNA delivery to plant cells.' *Plant Cell Reports* 11(7), 323–328

Finnegan, E. J., and E. S. Dennis (1993) 'Isolation and identification by sequence homology of a putative cytosine methyltransferase from Arabidopsis thaliana.' *Nucleic Acids Research* 21(10), 2383–2388

Finnegan, E. J., R. K. Genger, W. J. Peacock, and E. S. Dennis (1998) 'DNA methylation in plants.' *Annual Review of Plant Physiology and Plant Molecular Biology* 49, 223–247

Ford, R. H. (2000) 'Inheritance of kernel color in corn: Explanations and investigations.' *The American Biology Teacher* 62(3), 181–188

Frommer, M., L. E. McDonald, D. S. Millar, C. M. Collis, F. Watt, G. W. Grigg, P. L. Molloy, and C. L. Paul (1992) 'A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands.' *PNAS* 89(5), 1827–1831

Fujimoto, R., T. Sasaki, H. Kudoh, J. M. Taylor, T. Kakutani, and E. S. Dennis (2011) 'Epigenetic variation in the FWA gene within the genus Arabidopsis.' *The Plant Journal* 66(5), 831–843

Fujimoto, R., Y. Kinoshita, A. Kawabe, T. Kinoshita, K. Takashima, M. Nordborg, M. E. Nasrallah, K. K. Shimizu, H. Kudoh, and T. Kakutani (2008) 'Evolution and control of imprinted FWA genes in the genus Arabidopsis.' *PLoS Genetics* 4(4), e1000048

Gao, S., R. Babu, Y. Lu, C. Martinez, Z. Hao, A. F. Krivanek, J. Wang, T. Rong, J. Crouch, and Y. Xu (2011) 'Revisiting the hetero-fertilization phenomenon in maize.' *PLoS One* 6(1), e16101

Gaudin, V., M. Libault, S. Pouteau, T. Juul, G. Zhao, D. Lefebvre, and O. Grandjean (2001) 'Mutations in LIKE HETEROCHROMATIN PROTEIN 1 affect flowering time and plant architecture in Arabidopsis.' *Development* 128(23), 4847–4858

Gehring, M., J. H. Huh, T. F. Hsieh, J. Penterman, Y. Choi, J. J. Harada, R. B. Goldberg, and R. L. Fischer (2006) 'DEMETER DNA glycosylase establishes MEDEA polycomb gene self-imprinting by allele-specific demethylation.' *Cell* 124(3), 495–506

Gehring, M., K. L. Bubb, and S. Henikoff (2009) 'Extensive demethylation of repetitive elements during seed development underlies gene imprinting.' *Science* 324(5933), 1447–1451

Goff, S. A., T. M. Klein, B. A. Roth, M. E. Fromm, K. C. Cone, J. P. Radicella, and V. L. Chandler (1990) 'Transactivation of anthocyanin biosynthetic genes following transfer of B regulatory genes into maize tissues.' *EMBO Journal* 9(8), 2517–2522

Gong, Z., T. Morales-Ruiz, R. R. Ariza, T. Roldán-Arjona, L. David, and JK. Zhu (2002) 'ROS1, a repressor of transcriptional gene silencing in Arabidopsis, encodes a DNA glycosylase/lyase.' *Cell* 111(6), 803–814

Greaves, I. K., M. Groszmann, H. Ying, J. M. Taylor, W. J. Peacock, and E. S. Dennis (2012) 'Trans chromosomal methylation in Arabidopsis hybrids.' *PNAS* 109(9), 3570–3575

Greenberg, M. V., I. Ausin, S. W. Chan, S. J. Cokus, J. T. Cuperus, S. Feng, J. A. Law, C. Chu, M. Pellegrini, and J. C. Carrington (2011) 'Identification of genes required for de novo DNA methylation in Arabidopsis.' *Epigenetics* 6(3), 344–354

Gruenbaum, Y., T. Naveh-Many, H. Cedar, and A. Razin (1981) 'Sequence specificity of methylation in higher plant DNA.' *Nature* 292(5826), 860–862

Guignard, L. (1899) 'Sur les antherozoides et la double copulation sexualle chex les sur les antherzoides et la doube copulation sexualle chex les vegetaux angiosperms.' *Rev Gen Bot* 11, 129–135

Gutiérrez-Marcos, J. F., L. M. Costa, M. Dal Prà, S. Scholten, E. Kranz, P. Perez, and H. G. Dickinson (2006) 'Epigenetic asymmetry of imprinted genes in plant gametes.' *Nature Genetics* 38(8), 876–878

Gutierrez-Marcos, J. F., P. D. Pennington, L. M. Costa, and H. G. Dickinson (2003) 'Imprinting in the endosperm: a possible role in preventing wide hybridization.' *Philisophical Transactions of the Royal Society B: Biological* 358(1434), 1105–1111

Haberer, G., S. Young, A. K. Bharti, H. Gundlach, C. Raymond, G. Fuks, E. Butler, R. A. Wing, S. Rounsley, and B. Birren (2005) 'Structure and architecture of the maize genome.' *Plant Physiology* 139(4), 1612–1624

Haig, D., and M. Westoby (1989) 'Parent-specific gene expression and the triploid endosperm.' *American Naturalist* 123(1), 147–155

Halbhuber, K. J., and K. König (2003) 'Modern laser scanning microscopy in biology, biotechnology and medicine.' *Annals of Anatomy* 185(1), 1–20

Hale, C. J., J. L. Stonaker, S. M. Gross, and J. B. Hollick (2007) 'A novel Snf2 protein maintains trans-generational regulatory states established by paramutation in maize.' *PLoS Biology* 5(10), 2156–2165

Hale, C. J., K. F. Erhard Jr, D. Lisch, and J. B. Hollick (2009) 'Production and processing of siRNA precursor transcripts from the highly repetitive maize genome.' *PLoS Genetics* 5(8), e1000598

Hamilton, A. J., and D. C. Baulcombe (1999) 'A species of small antisense RNA in posttranscriptional gene silencing in plants.' *Science* 286(5441), 950–952

Hammond, S. M., E. Bernstein, D. Beach, and G. J. Hannon (2000) 'An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells.' *Nature* 404(6775), 293–296

Hansen, K. H., A. P. Bracken, D. Pasini, N. Dietrich, S. S. Gehani, A. Monrad, J. Rappsilber, M. Lerdrup, and K. Helin (2008) 'A model for transmission of the H3K27me3 epigenetic mark.' *Nature Cell Biology* 10(11), 1291–1300

Haring, M., R. Bader, M. Louwers, A. Schwabe, R. van Driel, and M. Stam (2010) 'The role of DNA methylation, nucleosome occupancy and histone modifications in paramutation.' *The Plant Journal* 63(3), 366–378

Hart, C. M., B. Fischer, JM. Neuhaus, and F. Meins Jr (1992) 'Regulated inactivation of homologous gene expression in transgenic nicotiana sylvestris plants containing a defense-related tobacco chitinase gene.' *Molecular and General Genetics* 235(2-3), 179–188

Haun, W. J., and N. M. Springer (2008) 'Maternal and paternal alleles exhibit differential histone methylation and acetylation at maize imprinted genes.' *The Plant Journal* 56(6), 903–912

He, XJ., T. Chen, and JK Zhu (2011) 'Regulation and function of DNA methylation in plants and animals.' *Cell Research* 21(3), 442–465

He, XJ., YF. Hsu, S. Zhu, A. T. Wierzbicki, O. Pontes, C. S. Pikaard, HL. Liu, CS. Wang, H. Jin, and JK. Zhu (2009) 'An effector of RNA-directed DNA methylation in Arabidopsis is an ARGONAUTE 4-and RNA-binding protein.' *Cell* 137(3), 498–508

Henderson, I. R., X. Zhang, C. Lu, L. Johnson, B. C. Meyers, P. J. Green, and S. E. Jacobsen (2006) 'Dissecting Arabidopsis thaliana DICER function in small RNA processing, gene silencing and DNA methylation patterning.' *Nature Genetics* 38(6), 721–725

Henikoff, S. (1997) 'Nuclear organization and gene expression: homologous pairing and long-range interactions.' *Current Opin of Cellular Biology* 9(3), 388–395

Hollick, J. B. (2012) 'Paramutation: a trans-homolog interaction affecting heritable gene regulation.' *Current Opinion in Plant Biology* 15(5), 536–543

Hollick, J. B., and V. L. Chandler (2001) 'Genetic factors required to maintain repression of a paramutagenic maize pl1 allele.' *Genetics* 157(1), 369–378

Holliday, R., and J. E. Pugh (1975) 'DNA modification mechanisms and gene activity during development.' *Science* 187(4173), 226–232

Howell, M. D., N. Fahlgren, E. J. Chapman, J. S. Cumbie, C. M. Sullivan, S. A. Givan, K. D. Kasschau, and J. C. Carrington (2007) 'Genome-wide analysis of the RNA-DEPENDENT RNA POLYMERASE6/dicer-like4 pathway in Arabidopsis reveals dependency on miRNA-and tasiRNA-directed targeting.' *The Plant Cell* 19(3), 926–942

Hsieh, T. F., C. A. Ibarra, P. Silva, A. Zemach, L. Eshed-Williams, R. L. Fischer, and D. Zilberman (2009) 'Genome-wide demethylation of Arabidopsis endosperm.' *Science* 324(5933), 1451–1454

Huang, L., A. ME. Jones, I. Searle, K. Patel, H. Vogler, N. C. Hubner, and D. C. Baulcombe (2008) 'An atypical RNA polymerase involved in RNA silencing shares small subunits with RNA polymerase II.' *Nature Structural and Molecular Biology* 16(1), 91–93

Huh, J. H., M. J. Bauer, T. F. Hsieh, and R. L. Fischer (2008) 'Cellular programming of plant gene imprinting.' *Cell* 132(5), 735–744

Hunter, M. P., N. Ismail, X. Zhang, B. D. Aguda, E. J. Lee, L. Yu, T. Xiao, J. Schafer, M. L. Lee, T. D. Schmittgen, S. P. Nana-Sinkam, D. Jarjoura, and C. B. Marsh (2008) 'Detection of microRNA expression in human peripheral blood microvesicles.' *PLoS One* 3(11), e3964

Ibarra, C. A., X. Feng, V. K. Schoft, TF. Hsieh, R. Uzawa, J. A. Rodrigues, A. Zemach, N. Chumak, A. Machlicova, and T. Nishimura (2012) 'Active DNA demethylation in plant companion cells reinforces transposon methylation in gametes.' *Science* 337(6100), 1360–1364

Ingelbrecht, I., H. Van Houdt, M. Van Montagu, and A. Depicker (1994) 'Posttranscriptional silencing of reporter transgenes in tobacco correlates with DNA methylation.' *PNAS* 91(22), 10502–10506

Ingham, D. J., S. Beer, S. Money, and G. Hansen (2001) 'Quantitative real-time PCR assay for determining transgene copy number in transformed plants.' *Biotechniques* 31(1), 132–134

Inouye, S., and F. I. Tsuji (1994) 'Aequorea green fluorescent protein. expression of the gene and fluorescence characteristics of the recombinant protein.' *FEBS Letters* 341(2-3), 277–280

Jackson, J. P., A. M. Lindroth, X. Cao, and S. E. Jacobsen (2002) 'Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase.' *Nature* 416(6880), 556–560

Jahnke, S., and S. Scholten (2009) 'Epigenetic resetting of a gene imprinted in plant embryos.' *Current Biology* 19(19), 1677–1681

Jeddeloh, J. A., T. L. Stokes, and E. J. Richards (1999) 'Maintenance of genomic methylation requires a SWI2/SNF2-like protein.' *Nature Genetics* 22(1), 94–97

Jia, Y., D. R. Lisch, K. Ohtsu, M. J. Scanlon, D. Nettleton, and P. S. Schnable (2009) 'Loss of RNA-dependent RNA polymerase 2 (RDR2) function causes widespread and unexpected changes in the expression of transposons, genes, and 24-nt small RNAs.' *PLoS Genetics* 5(11), e1000737

Jiang, L., J. Zhang, JJ. Wang, L. Wang, L. Zhang, G. Li, X. Yang, X. Ma, X. Sun, and J. Cai (2013) 'Sperm, but not oocyte, DNA methylome is inherited by zebrafish early embryos.' *Cell* 153(4), 773–784

Johnson, L. M., M. Bostick, X. Zhang, E. Kraft, I. Henderson, J. Callis, and S. E. Jacobsen (2007) 'The SRA methyl-cytosine-binding domain links DNA and histone methylation.' *Current Biology* 17(4), 379–384

Jose, A. M., G. A. Garcia, and C. P. Hunter (2011) 'Two classes of silencing RNAs move between Caenorhabditis elegans tissues.' *Nature Structural and Molecular Biology* 18(11), 1184–1188

Joshi, A. A., and K. Struhl (2005) 'Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to pol II elongation.' *Molecular Cell* 20(6), 971

Kallioniemi, A., T. Visakorpi, R. Karhu, D. Pinkel, and O. P. Kallioniemi (1996) 'Gene copy number analysis by fluorescence in situ hybridization and comparative genomic hybridization.' *Methods* 9(1), 113–121

Kanazawa, A., M. O'Dell, and R. P. Hellens (2007a) 'The binding of nuclear factors to the as-1 element in the CaMV 35S promoter is affected by cytosine methylation in vitro.' *Plant Biology* 9(3), 435–441

_ (2007b) 'Epigenetic inactivation of chalcone synthase-A transgene transcription in petunia leads to a reversion of the post-transcriptional gene silencing phenotype.' *Plant and Cell Physiology* 48(4), 638–647

Kankel, M. W., D. E. Ramsey, T. L. Stokes, S. K. Flowers, J. R. Haag, J. A. Jeddeloh, N. C. Riddle, M. L. Verbsky, and E. J. Richards (2003) 'Arabidopsis MET1 cytosine methyltransferase mutants.' *Genetics* 163(3), 1109–1122

Kass, S. U., N. Landsberger, and A. P. Wolffe (1997) 'DNA methylation directs a time-dependent repression of transcription initiation.' *Current Biology* 7(3), 157–165

Kazazian, H. H. (2004) 'Mobile elements: drivers of genome evolution.' *Science* 303(5664), 1626–1632

Kermicle, J. L. (1970) 'Dependence of the R-mottled aleurone phenotype in maize on mode of sexual transmission.' *Genetics* 66(1), 69–85

Khaitová, L. C., M. Fojtová, K. Křížová, J. Lunerová, J. Fulneček, A. Depicker, and A. Kovařík (2011) 'Paramutation of tobacco transgenes by small RNA-mediated transcriptional gene silencing.' *Epigenetics* 6(5), 650–660

Kilby, N. J., HM. O. Leyser, and I. J. Furner (1992) 'Promoter methylation and progressive transgene inactivation in Arabidopsis.' *Plant Molecular Biology* 20(1), 103–112

Kinoshita, T., A. Miura, Y. Choi, Y. Kinoshita, X. Cao, S. E. Jacobsen, R. L. Fischer, and T. Kakutani (2004) 'One-way control of FWA imprinting in Arabidopsis endosperm by DNA methylation.' *Science* 303(5657), 521–523

Klein, T. M., and S. Fitzpatrick-Mcelligott (1993) 'Particle bombardment: a universal approach for gene transfer to cells and tissues.' *Current Opinion in Biotechnology* 4(5), 583–590

Klopfenstein, T. J., G. E. Erickson, and L. L. Berger (2012) 'Maize is a critically important source of food, feed, energy and forage in the USA.' *Field Crops Research*

Köhler, C., and I. Weinhofer-Molisch (2009) 'Mechanisms and evolution of genomic imprinting in plants.' *Heredity* 105(1), 57–63

Koukalová, B., I. Votruba, M. Fojtová, A. Holỳ, and A. Kovaŕík (2002) 'Hypomethylation of CNG targets induced with dihydroxypropyladenine is rapidly reversed in the course of mitotic cell division in tobacco.' *Theoretical and Applied Genetics* 105(5), 796–801

Kouzarides, T. (2007) 'Chromatin modifications and their function.' *Cell* 128(4), 693–705

Krichevsky, A. M., K. S. King, C. P. Donahue, K. Khrapko, and K. S. Kosik (2003) 'A microRNA array reveals extensive regulation of microRNAs during brain development.' *RNA* 9(10), 1274–1281

Lachner, M., D. O'Carroll, S. Rea, K. Mechtler, and T. Jenuwein (2001) 'Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins.' *Nature* 410(6824), 116–120

Lam, E., P. N. Benfey, P. M. Gilmartin, R. X. Fang, and N. H. Chua (1989) 'Site-specific mutations alter in vitro factor binding and change promoter expression pattern in transgenic plants.' *PNAS* 86(20), 7890–7894

Langmead, B., C. Trapnell, M. Pop, and S. L. Salzberg (2009) 'Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.' *Genome Biology* 10(3), R25

Larramendy, M. L., W. el Rifai, A. Kokkola, P. Puolakkainen, O. Monni, R. Salovaara, M. Aarnio, and S. Knuutila (1998) 'Comparative genomic hybridization reveals differences in DNA copy number changes between sporadic gastric carcinomas and gastric carcinomas from patients with hereditary nonpolyposis colorectal cancer.' *Cancer Genetics and Cytogenetics* 106(1), 62–65

Lauria, M., M. Rupe, M. Guo, E. Kranz, R. Pirona, A. Viotti, and G. Lund (2004) 'Extensive maternal DNA hypomethylation in the endosperm of Zea mays.' *The Plant Cell* 16(2), 510–522

Law, J. A., A. A. Vashisht, J. A. Wohlschlegel, and S. E. Jacobsen (2011) 'SHH1, a homeodomain protein required for DNA methylation, as well as RDR2, RDM4, and chromatin remodeling factors, associate with RNA polymerase IV.' *PLoS Genetics* 7(7), e1002195

Law, J. A., and S. E. Jacobsen (2010) 'Establishing, maintaining and modifying DNA methylation patterns in plants and animals.' *Nature Reviews Genetics* 11(3), 204–220

Law, J. A., J. Du, C. J. Hale, S. Feng, K. Krajewski, AM. S. Palanca, B. D. Strahl, D. J. Patel, and S. E. Jacobsen (2013) 'Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1.' *Nature* 498(7454), 385–389

Li, E., T. H. Bestor, and R. Jaenisch (1992) 'Targeted mutation of the DNA methyltransferase gene results in embryonic lethality.' *Cell* 69(6), 915–926

Li, H., M. Freeling, and D. Lisch (2010) 'Epigenetic reprogramming during vegetative phase change in maize.' *PNAS* 107(51), 22184–22189

Lin, BY. (1984) 'Ploidy barrier to endosperm development in maize.' *Genetics* 107(1), 103–115

Lindbo, J. A., L. Silva-Rosales, W. M. Proebsting, and W. G. Dougherty (1993) 'Induction of a highly specific antiviral state in transgenic plants: implications for regulation of gene expression and virus resistance.' *The Plant Cell* 5(12), 1749–1759

Linder, P., and G. W. Owttrim (2009) 'Plant RNA helicases: linking aberrant and silencing RNA.' *Trends in Plant Science* 14(6), 344–352

Lindroth, A. M., D. Shultis, Z. Jasencakova, J. Fuchs, L. Johnson, D. Schubert, D. Patnaik, S. Pradhan, J. Goodrich, and I. Schubert (2004) 'Dual histone H3 methylation marks at lysines 9 and 27 required for interaction with CHROMOMETHYLASE3.' *EMBO Journal* 23(21), 4146–4155

Lindroth, A. M., X. Cao, J. P. Jackson, D. Zilberman, C. M. McCallum, S. Henikoff, and S. E. Jacobsen (2001) 'Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation.' *Science* 292(5524), 2077–2080

Lisch, D. (2009) 'Epigenetic regulation of transposable elements in plants.' *Annual Review of Plant Biology* 60, 43–66

Lisch, D., C. C. Carey, J. E. Dorweiler, and V. L. Chandler (2002) 'A mutation that prevents paramutation in maize also reverses Mutator transposon methylation and silencing.' *PNAS* 99(9), 6130–6135

Lister, R., R. C. O'Malley, J. Tonti-Filippini, B. D. Gregory, C. C. Berry, A. H. Millar, and J. R. Ecker (2008) 'Highly integrated single-base resolution maps of the epigenome in Arabidopsis.' *Cell* 133(3), 523–536

Liu, C., F. Lu, X. Cui, and X. Cao (2010) 'Histone methylation in higher plants.' *Annual Review of Plant Biology* 61, 395–420

Liu, J., Y. He, R. Amasino, and X. Chen (2004) 'siRNAs targeting an intronic transposon in the regulation of natural flowering behavior in Arabidopsis.' *Genes and Development* 18(23), 2873–2878

Livak, K. J., and T. D. Schmittgen (2001) 'Analysis of relative gene expression data using real-time quantitative PCR and the 2(-DeltaDelta C(T)) method.' *Methods* 25(4), 402–408

Livet, J., T. A. Weissman, H. Kang, R. W. Draft, J. Lu, R. A. Bennis, J. R. Sanes, and J. W. Lichtman (2007) 'Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system.' *Nature* 450(7166), 56–62

Lu, C., S. S. Tej, S. Luo, C. D. Haudenschild, B. C. Meyers, and P. J. Green (2005) 'Elucidation of the small RNA component of the transcriptome.' *Science* 309(5740), 1567–1569

Lu, X., D. Chen, D. Shu, Z. Zhang, W. Wang, C. Klukas, L. Chen, Y. Fan, M. Chen, and C. Zhang (2013) 'The differential transcription network between embryo and endosperm in the early developing maize seed.' *Plant Physiology* 162(1), 440–455

Lucito, R., J. West, A. Reiner, J. Alexander, D. Esposito, B. Mishra, S. Powers, L. Norton, and M. Wigler (2000) 'Detecting gene copy number fluctuations in tumor cells by microarray analysis of genomic representations.' *Genome Research* 10(11), 1726–1736

Luger, K., A. W. Mäder, R. K. Richmond, D. F. Sargent, and T. J. Richmond (1997) 'Crystal structure of the nucleosome core particle at 2.8 Å resolution.' *Nature* 389(6648), 251–260

Lund, G., P. Ciceri, and A. Viotti (1995) 'Maternal-specific demethylation and expression of specific alleles of zein genes in the endosperm of Zea mays L.' *The Plant Journal* 8(4), 571–581

Madzima, T. F., E. S. Mills, J. M. Gardiner, and K. M. McGinnis (2011) 'Identification of epigenetic regulators of a transcriptionally silenced transgene in maize.' *G3 (Bethesda)* 1(1), 75–83

Makarevich, G., C. BR. Villar, A. Erilova, and C. Köhler (2008) 'Mechanism of PHERES1 imprinting in Arabidopsis.' *Journal of Cell Science* 121(6), 906–912

Malagnac, F., L. Bartee, and J. Bender (2002) 'An Arabidopsis SET domain protein required for maintenance but not establishment of DNA methylation.' *EMBO Journal* 21(24), 6842–6852

Mallory, A. C., and H. Vaucheret (2006) 'Functions of microRNAs and related small RNAs in plants.' *Nature Genetics* 38, S31–S36

Marenkova, T. V., D. B. Loginova, and E. V. Deineko (2012) 'Mosaic patterns of transgene expression in plants.' *Russian Journal of Genetics* 48(3), 249–260

Margueron, R., and D. Reinberg (2010) 'Chromatin structure and the inheritance of epigenetic information.' *Nature Reviews Genetics* 11(4), 285–296

Maruyama, D., Y. Hamamura, H. Takeuchi, D. Susaki, M. Nishimaki, D. Kurihara, R. D. Kasahara, and T. Higashiyama (2013) 'Independent control by each female gamete prevents the attraction of multiple pollen tubes.' *Developmental Cell* 25(3), 317–323

Mathieu, O., A. V. Probst, and J. Paszkowski (2005) 'Distinct regulation of histone H3 methylation at lysines 27 and 9 by CpG methylation in Arabidopsis.' *EMBO Journal* 24(15), 2783–2791

Matzke, A. J., and M. A. Matzke (1998) 'Position effects and epigenetic silencing of plant transgenes.' *Current Opinion in Plant Biology* 1(2), 142–148

Matzke, M. A., A. J. Matzke, and W. B. Eggleston (1996) 'Paramutation and transgene silencing: a common response to invasive DNA?' *Trends in Plant Science* 1(11), 382–388

Matzke, M. A., M. F. Mette, and A. J. Matzke (2000) 'Transgene silencing by the host genome defense: implications for the evolution of epigenetic control mechanisms in plants and vertebrates.' *Plant Molecular Biology* 43(2-3), 401–415

Matzke, M. A., M. Primig, J. Trnovsky, and A. J. Matzke (1989) 'Reversible methylation and inactivation of marker genes in sequentially transformed tobacco plants.' *EMBO Journal* 8(3), 643–649

Matzke, M. A., T. Kanno, L. Daxinger, B. Huettel, and A. J. Matzke (2009) 'RNA-mediated chromatin-based silencing in plants.' *Current Opinion in Cell Biology* 21(3), 367–376

McClintock, B. (1951) 'Chromosome organization and genic expression.' In 'Cold Spring Harbor Symposia on Quantitative Biology' number 16, 13-47 Cold Spring Harbor Laboratory Press

McCue, A. D., S. Nuthikattu, S. H. Reeder, and R. K. Slotkin (2012) 'Gene expression and stress response mediated by the epigenetic regulation of a transposable element small RNA.' *PLoS Genetics* 8(2), e1002474

McGinnis, K. M., C. Springer, Y. Lin, C. C. Carey, and V. L. Chandler (2006) 'Transcriptionally silenced transgenes in maize are activated by three mutations defective in paramutation.' *Genetics* 173(3), 1637–1647

McGrath, J., and D. Solter (1984) 'Completion of mouse embryogenesis requires both the maternal and paternal genomes.' *Cell* 37(1), 179–183

Melnyk, C. W., A. Molnar, A. Bassett, and D. C. Baulcombe (2011) 'Mobile 24 nt small RNAs direct transcriptional gene silencing in the root meristems of Arabidopsis thaliana.' *Current Biology* 21(19), 1678–1683

Mette, M. F., W. Aufsatz, J. van der Winden, M. A. Matzke, and A. J. Matzke (2000) 'Transcriptional silencing and promoter methylation triggered by double-stranded RNA.' *EMBO Journal* 19(19), 5194–5201

Meyer, P., and H. Saedler (1996) 'Homology-dependent gene silencing in plants.' *Annual Reviews of Plant Physiology and Plant Molecular Biology* 47, 23–48

Meyer, P., I. Heidmann, and I. Niedenhof (1993) 'Differences in DNA-methylation are associated with a paramutation phenomenon in transgenic petunia.' *The Plant Journal* 4(1), 89–100

Meyer, P., I. Niedenhof, and M. ten Lohuis (1994) 'Evidence for cytosine methylation of non-symmetrical sequences in transgenic Petunia hybrida.' *EMBO Journal* 13(9), 2084–2088

Meyers, B. C., S. S. Tej, T. H. Vu, C. D. Haudenschild, V. Agrawal, S. B. Edberg, H. Ghazal, and S. Decola (2004) 'The use of MPSS for whole-genome transcriptional analysis in Arabidopsis.' *Genome Research* 14(8), 1641–1653

Minksy, M. (1988) 'Memoir on inventing the confocal scanning microscope.' *Scanning* 10(4), 128–138

Mitsuhara, I., N. Shirasawa-Seo, T. Iwai, S. Nakamura, R. Honkura, and Y. Ohashi (2002) 'Release from post-transcriptional gene silencing by cell proliferation in transgenic tobacco plants: possible mechanism for noninheritance of the silencing.' *Genetics* 160(1), 343–352

Mittelsten, S. O., J. Paszkowski, and I. Potrykus (1991) 'Reversible inactivation of a transgene in Arabidopsis thaliana.' *Molecular Genetics* 228(1-2), 104–112

Molnar, A., C. W. Melnyk, A. Bassett, T. J. Hardcastle, R. Dunn, and D. C. Baulcombe (2010) 'Small silencing RNAs in plants are mobile and direct epigenetic modification in recipient cells.' *Science* 328(5980), 872–875

Moore, T. (2001) 'Southern analysis using methyl-sensitive restriction enzymes.' *Methods in Molecular Biology* 181, 193–203

Morgan, H. D., F. Santos, K. Green, W. Dean, and W. Reik (2005) 'Epigenetic reprogramming in mammals.' *Human Molecular Genetics* 14, R47–R58

Morino, K., OA. Olsen, and K. Shimamoto (1999) 'Silencing of an aleurone-specific gene in transgenic rice is caused by a rearranged transgene.' *The Plant Journal* 17(3), 275–285

Mosher, R. A., and C. W. Melnyk (2010) 'siRNAs and DNA methylation: seedy epigenetics.' *Trends in Plant Science* 15(4), 204–210

Mosher, R. A., F. Schwach, D. Studholme, and D. C. Baulcombe (2008) 'PolIVb influences RNA-directed DNA methylation independently of its role in siRNA biogenesis.' *PNAS* 105(8), 3145–3150

Nagaya, S., K. Kato, Y. Ninomiya, R. Horie, M. Sekine, K. Yoshida, and A. Shinmyo (2005) 'Expression of randomly integrated single complete copy transgenes does not vary in Arabidopsis thaliana.' *Plant Cell Physiology* 46(3), 438–444

Nan, X., R. R. Meehan, and A. Bird (1993) 'Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2.' *Nucleic Acids Research* 21(21), 4886–4892

Napoli, C., C. Lemieux, and R. Jorgensen (1990) 'Introduction of a chimeric chalcone synthase gene into Petunia results in reversible co-suppression of homologous genes in trans.' *Plant Cell* 2(4), 279–289

Nawaschin, S. (1898) 'Resultate einer revision der befruchtungsvorgange bei lilium martagon und fritillaria tenella.' *Lilium martagon and Fritillia tenella* 9(4), 377–382

Nobuta, K., C. Lu, R. Shrivastava, M. Pillay, E. De Paoli, M. Accerbi, M. Arteaga-Vazquez, L. Sidorenko, D. H. Jeong, Y. Yen, P. J. Green, V. L. Chandler, and B. C. Meyers (2008) 'Distinct size distribution of endogeneous siRNAs in maize: Evidence from deep sequencing in the mop1-1 mutant.' *PNAS* 105(39), 14958–14963

Nodine, M. D., and D. P. Bartel (2012) 'Maternal and paternal genomes contribute equally to the transcriptome of early plant embryos.' *Nature* 482(7383), 94–97

Norton, V. G., B. S. Imai, P. Yau, and E. M. Bradbury (1989) 'Histone acetylation reduces nucleosome core particle linking number change.' *Cell* 57(3), 449–457

Numa, H., JM. Kim, A. Matsui, Y. Kurihara, T. Morosawa, J. Ishida, Y. Mochizuki, H. Kimura, K. Shinozaki, and T. Toyoda (2009) 'Transduction of RNA-directed DNA methylation signals to repressive histone marks in Arabidopsis thaliana.' *The EMBO Journal* 29(2), 352–362

Nuthikattu, S., A. D. McCue, K. Panda, D. Fultz, C. DeFraia, E. N. Thomas, and R. K. Slotkin (2013) 'The initiation of epigenetic silencing of active transposable elements is triggered by RDR6 and 21-22 nucleotide small interfering RNAs.' *Plant Physiology* 162(1), 116–131

Ohlsson, R., and C. Kanduri (2002) 'New twists on the epigenetics of CpG islands.' *Genome Research* 12(4), 525–526

Onodera, Y., J. R. Haag, T. Ream, P. C. Nunes, O. Pontes, and C. S. Pikaard (2005) 'Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation.' *Cell* 120(5), 613–622

Ortega-Galisteo, AP., T. Morales-Ruiz, R. R. Ariza, and T. Roldán-Arjona (2008) 'Arabidopsis DEMETER-LIKE proteins DML2 and DML3 are required for appropriate distribution of DNA methylation marks.' *Plant Molecular Biology* 67(6), 671–681

Ow, D. W., J. D. Jacobs, and S. H. Howell (1987) 'Functional regions of the cauliflower mosaic virus 35S RNA promoter determined by use of the firefly luciferase gene as a reporter of promoter activity.' *PNAS* 84(14), 4870–4874

Paddock, S. W. (2000) 'Principles and practices of laser scanning confocal microscopy.' *Molecular Biotechnology* 16(2), 127–149

Pan, YB., and P. A. Peterson (1988) 'Spontaneous activation of quiescent transposable elements during endosperm development in Zea mays.' *Genetics* 119(2), 457–464

Park, Y. D., I. Papp, E. A. Moscone, V. A. Iglesias, H. Vaucheret, A. J. Matzke, and M. A. Matzke (1996) 'Gene silencing mediated by promoter homology occurs at the level of transcription and results in meiotically heritable alterations in methylation and gene activity.' *The Plant Journal* 9(2), 183–194

Paszkowski, J., and S. A Whitham (2001) 'Gene silencing and DNA methylation processes.' *Current Opinion in Plant Biology* 4(2), 123–129

Peng, M., Y. Cui, YM. Bi, and S. J. Rothstein (2006) 'AtMBD9: a protein with a methyl-CpG-binding domain regulates flowering time and shoot branching in Arabidopsis.' *The Plant Journal* 46(2), 282–296

Petroni, K., E. Cominelli, G. Consonni, G. Gusmaroli, G. Gavazzi, and C. Tonelli (2000) 'The developmental expression of the maize regulatory gene hopi determines germination-dependent anthocyanin accumulation.' *Genetics* 155(1), 323–336

Pikaard, C. S., and S. Tucker (2009) 'RNA-silencing enzymes Pol IV and Pol V in maize: more than one flavor?' *PLoS Genetics* 5(11), e1000736

Pikaard, C. S., J. R. Haag, T. Ream, and A. T. Wierzbicki (2008) 'Roles of RNA polymerase IV in gene silencing.' *Trends in Plant Science* 13(7), 390–397

Pontier, D., C. Picart, F. Roudier, D. Garcia, S. Lahmy, J. Azevedo, E. Alart, M. Laudié, W. M. Karlowski, and R. Cooke (2012) 'NERD, a plant-specific GW protein, defines an additional RNAi-dependent chromatin-based pathway in Arabidopsis.' *Molecular Cell* 48(1), 121–132

Prasad, V., D. Semwogerere, and E. R. Weeks (2007) 'Confocal microscopy of colloids.' *Journal of Physics: Condensed Matter* 19(11), 113102

Prasher, D. C., V. K. Eckenrode, W. W. Ward, F. G. Prendergast, and M. J. Cormier (1992) 'Primary structure of the Aequorea victoria green-fluorescent protein.' *Gene* 111(2), 229–233

Prendergast, G. C., and E. B. Ziff (1991) 'Methylation-sensitive sequence-specific DNA binding by the c-Myc basic region.' *Science* 251(4990), 186–189

Pröls, F., and P. Meyer (1992) 'The methylation patterns of chromosomal integration regions influence gene activity of transferred DNA in Petunia hybrida.' *The Plant Journal* 2(4), 465–475

Qin, H., Y. Dong, and A. G. von Arnim (2003) 'Epigenetic interactions between Arabidopsis transgenes: characterization in light of transgene integration sites.' *Plant Molecular Biology* 52(1), 217–231

Raissig, M. T., C. Baroux, and U. Grossniklaus (2011) 'Regulation and flexibility of genomic imprinting during seed development.' *Plant Cell* 23(1), 16–26

Rangwala, S. H., R. Elumalai, C. Vanier, H. Ozkan, D. W. Galbraith, and E. J. Richards (2006) 'Meiotically stable natural epialleles of sadhu, a novel Arabidopsis retroposon.' *PLoS Genetics* 2(3), e36

Regulski, M., Z. Lu, J. Kendall, M. TA. Donoghue, J. Reinders, V. Llaca, S. Deschamps, A. Smith, D. Levy, and W. R. McCombie (2013) 'The maize methylome influences mRNA splice sites and reveals widespread paramutation-like switches guided by small RNA.' *Genome Research* 23, 1651–1662

Reik, W., A. Collick, M. L. Norris, S. C. Barton, and M. A. Surani (1987) 'Genomic imprinting determines methylation of parental alleles in transgenic mice.' *Nature* 328(6127), 248–251

Robertson, D. S. (1984) 'A study of heterofertilization in diverse lines of maize.' *Journal of Heredity* 75(6), 457–462

Rochester, D. E., J. A. Winer, and D. M. Shah (1986) 'The structure and expression of maize genes encoding the major heat shock protein, hsp70.' *EMBO Journal* 5(3), 451

Rodin, S. N., and A. D. Riggs (2003) 'Epigenetic silencing may aid evolution by gene duplication.' *Journal of Molecular Evolution* 56(6), 718–729

Rossi, V., S. Locatelli, C. Lanzanova, M. B. Boniotti, S. Varotto, A. Pipal, M. Goralik-Schramel, A. Lusser, C. Gatz, C. Gutierrez, and M. Motto (2003) 'A maize histone deacetylase and retinoblastoma-related protein physically interact and cooperate in repressing gene transcription.' *Plant Molecular Biology* 51(3), 401–413

Rozen, S., and H. Skaletsky (1999) 'Primer3 on the WWW for general users and for biologist programmers.' *Methods of Molecular Biology* 132, 365–386

Sarkar, K. R., and E. H. Coe (1971) 'Analysis of events leading to heterofertilization in maize.' *Journal of Heredity* 62(2), 118–120

Sarraf, S. A., and I. Stancheva (2004) 'Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly.' *Molecular Cell* 15(4), 595–605

Saze, H., and T. Kakutani (2007) 'Heritable epigenetic mutation of a transposon-flanked Arabidopsis gene due to lack of the chromatin-remodeling factor DDM1.' *EMBO Journal* 26(15), 3641–3652

Saze, H., K. Tsugane, T. Kanno, and T. Nishimura (2012) 'DNA methylation in plants: relationship to small RNAs and histone modifications, and functions in transposon inactivation.' *Plant and Cell Physiology* 53(5), 766–784

Saze, H., O. M. Scheid, and J. Paszkowski (2003) 'Maintenance of CpG methylation is essential for epigenetic inheritance during plant gametogenesis.' *Nature Genetics* 34(1), 65–69

Scheid, O. M., K. Afsar, and J. Paszkowski (1998) 'Release of epigenetic gene silencing by trans-acting mutations in Arabidopsis.' *PNAS* 95(2), 632–637

Schmittgen, T. D., J. Jiang, Q. Liu, and L. Yang (2004) 'A high-throughput method to monitor the expression of microRNA precursors.' *Nucleic Acids Research* 32(4), e43

Schnable, P. S., D. Ware, R. S. Fulton, J. C. Stein, F. Wei, S. Pasternak, C. Liang, J. Zhang, L. Fulton, T. A. Graves, P. Minx, A. D. Reily, L. Courtney, S. S. Kruchowski, C. Tomlinson, C. Strong, K. Delehaunty, C. Fronick, B. Courtney, S. M. Rock, E. Belter, F. Du, K. Kim, R. M. Abbott, M. Cotton, A. Levy, P. Marchetto, K. Ochoa, S. M. Jackson, B. Gillam, W. Chen, L. Yan, J. Higginbotham, M. Cardenas, J. Waligorski, E. Applebaum, L. Phelps, J. Falcone, K. Kanchi, T. Thane, A. Scimone, N. Thane, J. Henke, T. Wang, J. Ruppert, N. Shah, K. Rotter, J. Hodges, E. Ingenthron, M. Cordes, S. Kohlberg, J. Sgro, B. Delgado, K. Mead, A. Chinwalla, S. Leonard, K. Crouse, K. Collura, D. Kudrna, J. Currie, R. He, A. Angelova, S. Rajasekar,

T. Mueller, R. Lomeli, G. Scara, A. Ko, K. Delaney, M. Wissotski, G. Lopez, D. Campos, M. Braidotti, E. Ashley, W. Golser, H. Kim, S. Lee, J. Lin, Z. Dujmic, W. Kim, J. Talag, A. Zuccolo, C. Fan, A. Sebastian, M. Kramer, L. Spiegel, L. Nascimento, T. Zutavern, B. Miller, C. Ambroise, S. Muller, W. Spooner, A. Narechania, L. Ren, S. Wei, S. Kumari, B. Faga, M. J. Levy, L. McMahan, P. Van Buren, M. W. Vaughn, K. Ying, C. T. Yeh, S. J. Emrich, Y. Jia, A. Kalyanaraman, A. P. Hsia, W. B. Barbazuk, R. S. Baucom, T. P. Brutnell, N. C. Carpita, C. Chaparro, J. M. Chia, J. M. Deragon, J. C. Estill, Y. Fu, J. A. Jeddeloh, Y. Han, H. Lee, P. Li, D. R. Lisch, S. Liu, Z. Liu, D. H. Nagel, M. C. McCann, P. SanMiguel, A. M. Myers, D. Nettleton, J. Nguyen, B. W. Penning, L. Ponnala, K. L. Schneider, D. C. Schwartz, A. Sharma, C. Soderlund, N. M. Springer, Q. Sun, H. Wang, M. Waterman, R. Westerman, T. K. Wolfgruber, L. Yang, Y. Yu, L. Zhang, S. Zhou, Q. Zhu, J. L. Bennetzen, R. K. Dawe, J. Jiang, N. Jiang, G. G. Presting, S. R. Wessler, S. Aluru, R. A. Martienssen, S. W. Clifton, W. R. McCombie, R. A. Wing, and R. K. Wilson (2009) 'The B73 maize genome: complexity, diversity, and dynamics.' *Science* 326(5956), 1112–1115

Schotta, G., A. Ebert, R. Dorn, and G. Reuter (2003) 'Position-effect variegation and the genetic dissection of chromatin regulation in Drosophila.' *Seminars in cell developmental biology* 14(1), 67–75

Schubert, D., B. Lechtenberg, A. Forsbach, M. Gils, S. Bahadur, and R. Schmidt (2004) 'Silencing in Arabidopsis T-DNA transformants: the predominant role of a gene-specific RNA sensing mechanism versus position effects.' *The Plant Cell* 16(10), 2561–2572

Segal, E., and J. Widom (2009) 'From DNA sequence to transcriptional behaviour: a quantitative approach.' *Nature Reviews Genetics* 10(7), 443–456

Sekhon, R. S., and S. Chopra (2009) 'Progressive loss of DNA methylation releases epigenetic gene silencing from a tandemly repeated maize Myb gene.' *Genetics* 181(1), 81–91

Sekhon, R. S., P. H. Wang, L. Sidorenko, V. L. Chandler, and S. Chopra (2012) 'Maize unstable factor for orange1 is required for maintaining silencing associated with paramutation at the pericarp color1 and booster1 loci.' *PLoS Genetics* 8(10), e1002980

Sharif, J., M. Muto, S. Takebayashi, I. Suetake, A. Iwamatsu, T. A. Endo, J. Shinga, Y. Mizutani-Koseki, T. Toyoda, and K. Okamura (2007) 'The SRA protein Np95 mediates epigenetic inheritance by recruiting dnmt1 to methylated DNA.' *Nature* 450(7171), 908–912

Sheldon, C. C., M. J. Hills, C. Lister, C. Dean, E. S. Dennis, and W. J. Peacock (2008) 'Resetting of FLOWERING LOCUS C expression after epigenetic repression by vernalization.' *PNAS* 105(6), 2214–2219

Shimomura, O., F. H. Johnson, and Y. Saiga (1962) 'Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea.' *Journal of Cellular and Comparative Physiology* 59, 223–239

Sidorenko, L., J. E. Dorweiler, A. M. Cigan, M. Arteaga-Vazquez, M. Vyas, J. Kermicle, D. Jurcin, J. Brzeski, Y. Cai, and V. L. Chandler (2009) 'A dominant mutation in mediator of paramutation2, one of three second-largest subunits of a plant-specific RNA polymerase, disrupts multiple sirna silencing processes.' *PLoS Genetics* 5(11), e1000725

Sidorenko, L., Lyudmila V., and T. Peterson (2001) 'Transgene-induced silencing identifies sequences involved in the establishment of paramutation of the maize p1 gene.' *The Plant Cell* 13(2), 319–335

Siebert, P. D., A. Chenchik, D. E. Kellogg, K. A. Lukyanov, and S. A. Lukyanov (1995) 'An improved PCR method for walking in uncloned genomic DNA.' *Nucleic Acids Research* 23(6), 1087

Simon, S. A., and B. C. Meyers (2011) 'Small RNA-mediated epigenetic modifications in plants.' *Current Opinion in Plant Biology* 14(2), 148–155

Singh, J., M. Freeling, and D. Lisch (2008) 'A position effect on the heritability of epigenetic silencing.' *PLoS Genetics* 4(10), e1000216

Slotkin, R. K., M. Vaughn, F. Borges, M. Tanurdžić, J. D. Becker, J. A. Feijó, and R. A. Martienssen (2009) 'Epigenetic reprogramming and small RNA silencing of transposable elements in pollen.' *Cell* 136(3), 461–472

Song, P., C. Cai, M. Skokut, B. Kosegi, and J. Petolino (2002) 'Quantitative real-time PCR as a screening tool for estimating transgene copy number in WHISKERS™-derived transgenic maize.' *Plant Cell Reports* 20(10), 948–954

Soppe, W. J., S. E. Jacobsen, C. Alonso-Blanco, J. P. Jackson, T. Kakutani, M. Koornneef, and A. JM. Peeters (2000) 'The late flowering phenotype of fwa mutants is caused by gain-of-function epigenetic alleles of a homeodomain gene.' *Molecular Cell* 6(4), 791–802

Sprague, G. F. (1929) 'Hetero-fertilization in maize.' *Science* 69, 526–527

_ (1932) 'The nature and extent of hetero-fertilization in maize.' *Genetics* 17(3), 358

Stam, M., C. Belele, W. Ramakrishna, J. E. Dorweiler, J. L. Bennetzen, and V. L. Chandler (2002) 'The regulatory regions required for B' paramutation and expression are located far upstream of the maize b1 transcribed sequences.' *Genetics* 162(2), 917–930

Stam, M., J. NM. Mol, and J. M. Kooter (1997) 'Review article: the silence of genes in transgenic plants.' *Annals of Botany* 79(1), 3–12

Steward, N., T. Kusano, and H. Sano (2000) 'Expression of ZmMET1, a gene encoding a DNA methyltransferase from maize, is associated not only with DNA replication in actively proliferating cells, but also with altered DNA methylation status in cold-stressed quiescent cells.' *Nucleic Acids Research* 28(17), 3250–3259

Stokes, T. L., and E. J. Richards (2002) 'Induced instability of two Arabidopsis constitutive pathogen-response alleles.' *PNAS* 99(11), 7792–7796

Stonaker, J. L., J. P. Lim, K. F. Erhard Jr, and J. B. Hollick (2009) 'Diversity of Pol IV function is defined by mutations at the maize rmr7 locus.' *PLoS Genetics* 5(11), e1000706

Strahl, B. D., and C. D. Allis (2000) 'The language of covalent histone modifications.' *Nature* 403(6765), 41–45

Stroud, H., M. VC. Greenberg, S. Feng, Y. V. Bernatavichute, and S. E. Jacobsen (2013) 'Comprehensive analysis of silencing mutants reveals complex regulation of the Arabidopsis methylome.' *Cell* 152(1), 352–364

Surani, M. A. (2001) 'Reprogramming of genome function through epigenetic inheritance.' *Nature* 414(6859), 122–128

Surani, M. A., S. C. Barton, and M. L. Norris (1984) 'Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis.' *Nature* 308(5959), 548–550

Susan, J. C., J. Harrison, C. L. Paul, and M. Frommer (1994) 'High sensitivity mapping of methylated cytosines.' *Nucleic Acids Research* 22(15), 2990–2997

Suter, C. M., and D. I. Martin (2010) 'Paramutation: the tip of an epigenetic iceberg?' *Trends in Genetics* 26(1), 9–14

Sutherland, E., L. Coe, and E. A. Raleigh (1992) 'McrBC: a multisubunit GTP-dependent restriction endonuclease.' *Journal of Molecular Biology* 225(2), 327–348

Suzuki, T., T. Matsuzaki, H. Hagiwara, T. Aoki, and K. Takata (2007) 'Recent advances in fluorescent labeling techniques for fluorescence microscopy.' *Acta Histochem Cytochem* 40(5), 131–137

Tadros, W., and H. D. Lipshitz (2009) 'The maternal-to-zygotic transition: a play in two acts.' *Development* 136(18), 3033–3042

Takeda, S., and J. Paszkowski (2006) 'DNA methylation and epigenetic inheritance during plant gametogenesis.' *Chromosoma* 115(1), 27–35

Talbert, P. B., and S. Henikoff (2006) 'Spreading of silent chromatin: inaction at a distance.' *Nature Reviews Genetics* 7(10), 793–803

Tan, M., H. Luo, S. Lee, F. Jin, J. S. Yang, E. Montellier, T. Buchou, Z. Cheng, S. Rousseaux, and N. Rajagopal (2011) 'Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification.' *Cell* 146(6), 1016–1028

Tartof, K. D., C. Hobbs, and M. Jones (1984) 'A structural basis for variegating position effects.' *Cell* 37(3), 869–878

Tournier, B., M. Tabler, and K. Kalantidis (2006) 'Phloem flow strongly influences the systemic spread of silencing in GFP Nicotiana benthamiana plants.' *The Plant Journal* 47(3), 383–394

Tsien, R. Y. (1998) 'The green fluorescent protein.' *Annual Review of Biochemistry* 67(1), 509–544

Turck, F., F. Roudier, S. Farrona, ML. Martin-Magniette, E. Guillaume, N. Buisine, S. Gagnot, R. A. Martienssen, G. Coupland, and V. Colot (2007) 'Arabidopsis TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27.' *PLoS Genetics* 3(6), e86

Turner, B. M. (2000) 'Histone acetylation and an epigenetic code.' *Bioessays* 22(9), 836–845

Unnamalai, N., B. G. Kang, and W. S. Lee (2004) 'Cationic oligopeptide-mediated delivery of dsRNA for post-transcriptional gene silencing in plant cells.' *FEBS Letters* 566(1-3), 307–310

Urnov, F. D., and A. P. Wolffe (2001) 'Above and within the genome: epigenetics past and present.' *Journal of Mammary Gland Biolgy and Neoplasia* 6(2), 153–167

Vain, P., N. Keen, J. Murillo, C. Rathus, C. Nemes, and J. J. Finer (1993) 'Development of the particle inflow gun.' *Plant Cell, Tissue and Organ Culture* 33(3), 237–246

Vakoc, C. R., S. A. Mandat, B. A. Olenchock, and G. A. Blobel (2005) 'Histone H3 lysine 9 methylation and HP1γ are associated with transcription elongation through mammalian chromatin.' *Molecular Cell* 19(3), 381–391

Vaucheret, H. (1993) 'Identification of a general silencer for 19S and 35S promoters in a transgenic tobacco plant: 90 bp of homology in the promoter sequence are sufficient for trans-inactivation.' *Comptes rendus de l'Académie des sciences* 316(12), 1471–1483

Velten, J., C. Cakir, E. Youn, J. Chen, and C. I. Cazzonelli (2012) 'Transgene silencing and transgene-derived siRNA production in tobacco plants homozygous for an introduced atMYB90 construct.' *PLoS One* 7(2), e30141

Vicient, C. (2010) 'Transcriptional activity of transposable elements in maize.' *BMC Genomics* 11(1), 601

Voinnet, O. (2009) 'Origin, biogenesis, and activity of plant microRNAs.' *Cell* 136(4), 669–687

Waddington, C. H. (1942) 'The epigenotype.' *Endeavour* 1, 18–20

_ (1953) 'Genetic assimilation of an acquired character.' *Evolution* pp. 118–126

Wang, MB., C. Masuta, N. A. Smith, and H. Shimura (2012) 'RNA silencing and plant viral diseases.' *Molecular Plant-Microbe Interactions* 25(10), 1275–1285

Wang, Q., and H. K. Dooner (2006) 'Remarkable variation in maize genome structure inferred from haplotype diversity at the bz locus.' *PNAS* 103(47), 17644–17649

Wang, X., A. A. Elling, X. Li, N. Li, Z. Peng, G. He, H. Sun, Y. Qi, X. S. Liu, and X. W. Deng (2009a) 'Genome-wide and organ-specific landscapes of epigenetic modifications and their relationships to mRNA and small RNA transcriptomes in maize.' *The Plant Cell* 21(4), 1053–1069

_ (2009b) 'Genome-wide and organ-specific landscapes of epigenetic modifications and their relationships to mRNA and small RNA transcriptomes in maize.' *The Plant Cell* 21(4), 1053–1069

Wassenegger, M., S. Heimes, L. Riedel, and H. L. Sänger (1994) 'RNA-directed de novo methylation of genomic sequences in plants.' *Cell* 76(3), 567–576

Waterhouse, P. M., M. W. Graham, and MB. Wang (1998) 'Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA.' *PNAS* 95(23), 13959–13964

Waterhouse, P. M., MB. Wang, and T. Lough (2001) 'Gene silencing as an adaptive defence against viruses.' *Nature* 411(6839), 834–842

Waters, A. J., I. Makarevitch, S. R. Eichten, R. A. Swanson-Wagner, CT. Yeh, W. Xu, P. S. Schnable, M. W. Vaughn, M. Gehring, and N. M. Springer (2011) 'Parent-of-origin effects on gene expression and DNA methylation in the maize endosperm.' *The Plant Cell* 23(12), 4221–4233

Wierzbicki, A. T., J. R. Haag, and C. S. Pikaard (2008) 'Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes.' *Cell* 135(4), 635–648

Wierzbicki, A. T., R. Cocklin, A. Mayampurath, R. Lister, M. J. Rowley, B. D. Gregory, J. R. Ecker, H. Tang, and C. S. Pikaard (2012) 'Spatial and functional relationships among Pol V-associated loci, Pol IV-dependent siRNAs, and cytosine methylation in the Arabidopsis epigenome.' *Genes and Development* 26(16), 1825–1836

Wolffe, A. P., and D. Guschin (2000) 'Review: chromatin structural features and targets that regulate transcription.' *Journal of Structural Biology* 129(2-3), 102–122

Woo, H. R., O. Pontes, C. S. Pikaard, and E. J. Richards (2007) 'VIM1, a methylcytosine-binding protein required for centromeric heterochromatinization.' *Genes and Development* 21(3), 267–277

Woo, H. R., T. A. Dittmer, and E. J. Richards (2008) 'Three SRA-domain methylcytosine-binding proteins cooperate to maintain global CpG methylation and epigenetic silencing in Arabidopsis.' *PLoS Genetics* 4(8), e1000156

Woodhouse, M. R., M. Freeling, and D. Lisch (2006) 'The mop1 (mediator of paramutation1) mutant progressively reactivates one of the two genes encoded by the MuDR transposon in maize.' *Genetics* 172(1), 579–592

Wu, C., P. K. Diggle, and W. E. Friedman (2013) 'Kin recognition within a seed and the effect of genetic relatedness of an endosperm to its compatriot embryo on maize seed development.' *PNAS* 110(6), 2217–2222

Wu, S. C., and Y. Zhang (2010) 'Active DNA demethylation: many roads lead to rome.' *Nature Reviews Molecular Cell Biology* 11(9), 607–620

Xiao, C., D. P. Calado, G. Galler, T. H. Thai, H. C. Patterson, J. Wang, N. Rajewsky, T. P. Bender, and K. Rajewsky (2007) 'MiR-150 controls B cell differentiation by targeting the transcription factor c-Myb.' *Cell* 131(1), 146–159

Xu, Y., Y. Wang, H. Stroud, X. Gu, B. Sun, ES. Gan, KH. Ng, S. E. Jacobsen, Y. He, and T. Ito (2013) 'A matrix protein silences transposons and repeats through interaction with retinoblastoma-associated proteins.' *Current Biology* 23(4), 345–350

Xue, W., C. Ruprecht, N. Street, K. Hematy, C. Chang, W. B. Frommer, S. Persson, and T. Niittylä (2012) 'Paramutation-like interaction of T-DNA loci in Arabidopsis.' *PloS One* 7(12), e51651

Yang, L., J. Ding, C. Zhang, J. Jia, H. Weng, W. Liu, and D. Zhang (2005) 'Estimating the copy number of transgenes in transformed rice by real-time quantitative PCR.' *Plant Cell Reports* 23(10-11), 759–763

You, W., A. Tyczewska, M. Spencer, L. Daxinger, M. W. Schmid, U. Grossniklaus, S. A. Simon, B. C. Meyers, A. J. Matzke, and M. A Matzke (2012) 'Atypical DNA methylation of genes encoding cysteine-rich peptides in Arabidopsis thaliana.' *BMC Plant Biology* 12(1), 51

Yuan, J. S., J. Burris, N. R. Stewart, A. Mentewab, and C. N. Stewart (2007) 'Statistical tools for transgene copy number estimation based on real-time PCR.' *BMC Bioinformatics*

Yun, M., J. Wu, J. L. Workman, and B. Li (2011) 'Readers of histone modifications.' *Cell Research* 21(4), 564–578

Zemach, A., and G. Grafi (2003) 'Characterization of Arabidopsis thaliana methyl-CpG-binding domain (MBD) proteins.' *The Plant Journal* 34(5), 565–572

_ (2007) 'Methyl-CpG-binding domain proteins in plants: interpreters of DNA methylation.' *Trends in Plant Science* 12(2), 80–85

Zemach, A., I. E. McDaniel, P. Silva, and D. Zilberman (2010) 'Genome-wide evolutionary analysis of eukaryotic DNA methylation.' *Science* 328(5980), 916–919

Zhang, H., X. Deng, D. Miki, S. Cutler, H. La, Y. J. Hou, J. Oh, and J. K. Zhu (2012) 'Sulfamethazine suppresses epigenetic silencing in Arabidopsis by impairing folate synthesis.' *The Plant Cell* 24(3), 1230–1241

Zhang, H., ZY. Ma, L. Zeng, K. Tanaka, CJ. Zhang, J. Ma, G. Bai, P. Wang, SW. Zhang, and ZW. Liu (2013) 'DTF1 is a core component of RNA-directed DNA methylation and may assist in the recruitment of Pol IV.' *PNAS* 110(20), 8290–8295

Zhang, X., J. Yazaki, A. Sundaresan, S. Cokus, S. WL. Chan, H. Chen, I. R. Henderson, P. Shinn, M. Pellegrini, and S. E. Jacobsen (2006) 'Genome-wide high-resolution mapping and functional analysis of DNA methylation in Arabidopsis.' *Cell* 126(6), 1189–1201

Zhang, X., O. Clarenz, S. Cokus, Y. V. Bernatavichute, M. Pellegrini, J. Goodrich, and S. E. Jacobsen (2007) 'Whole-genome analysis of histone H3 lysine 27 trimethylation in Arabidopsis.' *PLoS Biology* 5(5), e129

Zhang, X., Y. V. Bernatavichute, S. Cokus, M. Pellegrini, and S. E. Jacobsen (2009) 'Genome-wide analysis of mono-, di-and trimethylation of histone H3 lysine 4 in Arabidopsis thaliana.' *Genome Biology* 10(6), R62

Zheng, B., Z. Wang, S. Li, B. Yu, JY. Liu, and X. Chen (2009) 'Intergenic transcription by RNA polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in Arabidopsis.' *Genes and Development* 23(24), 2850–2860

Zhong, SH., JZ. Liu, H. Jin, L. Lin, Q. Li, Y. Chen, YX. Yuan, ZY. Wang, H. Huang, and YZ. Qi (2013) 'Warm temperatures induce transgenerational epigenetic release of RNA silencing by inhibiting siRNA biogenesis in Arabidopsis.' *PNAS* 110(22), 9171–9176

Zhong, X., C. J. Hale, J. A. Law, L. M. Johnson, S. Feng, A. Tu, and S. E. Jacobsen (2012) 'DDR complex facilitates global association of RNA polymerase V to promoters and evolutionarily young transposons.' *Nature Structural and Molecular Biology* 19(9), 870–875

Zhou, Y., J. Zhang, H. Lin, G. Guo, and Y. Guo (2010) 'MORPHEUS'MOLECULE1 is required to prevent aberrant RNA transcriptional read-through in Arabidopsis.' *Plant Physiology* 154(3), 1272–1280

Zilberman, D., D. Coleman-Derr, T. Ballinger, and S. Henikoff (2008) 'Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks.' *Nature* 456(7218), 125–129

Zilberman, D., M. Gehring, R. K. Tran, T. Ballinger, and S. Henikoff (2007) 'Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription.' *Nature Genetics* 39(1), 61–69

7 Appendix

7.1 Media Recipes

LB Media

10g/L Bacto-tryptone

5g/L Yeast Extract

10g/L NaCl

pH 7.5 (NaOH)

in addition:

1.5% w/v agar

MS media

4.4g/L Murashige and Skoog + vitamins

30g/L Sucrose

0.9% w/v agar

pH 5.8 (KOH)

in addition:

2mg/L 2,4-Dichlorophenoxyacetic acid

100mg/L Casein Hydrolysate

SOC media

20g/L Bacto-tryptone

5g/L Yeast Extract

0.5g/L NaCl

2.5mM KCl

10mM $MgCl_2$

20mM Glucose

pH 7.0 (NaOH)

7.2 Appendices for Chapter 3

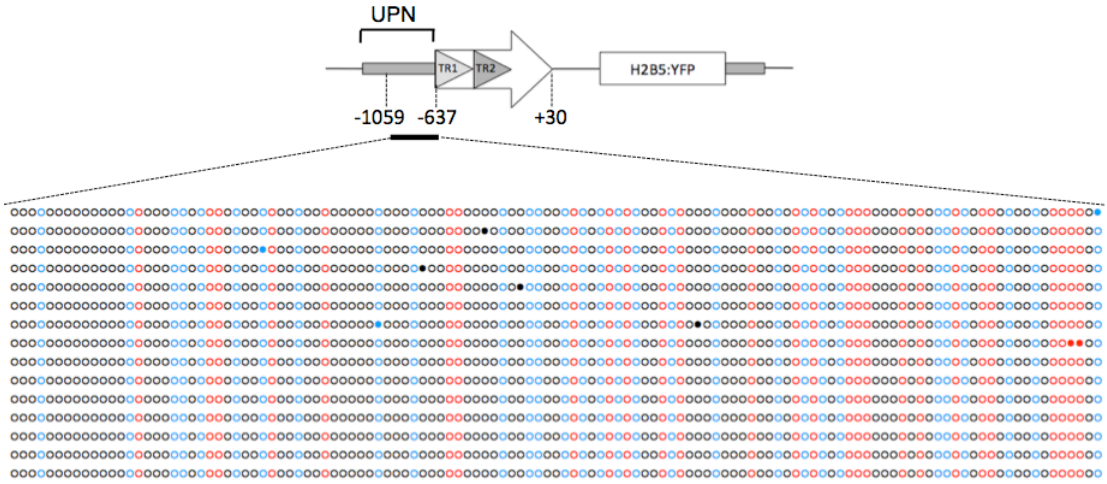


Figure 68: Methylation status of cytosines the UPN region of NYR-v in a WT background. Schematic diagram of NYR transgene and lollipop diagram of cytosine residues between positions -1059 to -637, displaying unmethylated cytosines (open circles), methylated cytosines (filled circles), site of asymmetric CHH methylation (black) and symmetric CG (red) or CHG methylation (blue) of 15 clones from bisulfite sequencing.

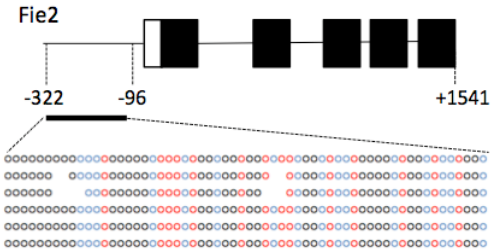


Figure 69: Methylation status of cytosines in *Fie2*. Schematic diagram of Fie2 (black boxes represent exons) region and lollipop diagram (positions -322 to -96) displaying unmethylated cytosines (open circles), methylated cytosines (filled circles), site of asymmetric CHH methylation (black) and symmetric CG (red) or CHG methylation (blue) of 6 clones from bisulfite sequencing. (Figure adapted from Gutierrez-Marcos et al. 2006)

Table 17: Root treatment of NYR-v transgenics with DNA methylation inhibitors

Treatment	Total number of seeds	Number NYR-v	Number NYR-a
untreated	20	20	0
40µM Zebularine	20	20	0
80µM Zebularine	20	20	0
100µM DHPA	20	20	0
200µM DHPA	20	20	0

Table 18: Seed treatment of NYR-v transgenics with DNA methylation inhibitors

Treatment	Total number of seeds	Total germination	Number NYR-v	Number NYR-a
untreated	30	28	12	0
40µM Zebularine	30	26	11	0
80µM Zebularine	30	26	15	0
100µM DHPA	30	29	16	0
200µM DHPA	30	27	15	0
100µM Sulfamethazine	30	29	10	0
200µM Sulfamethazine	30	29	17	0
5mM 5-Aza	30	28	12	0
10mM 5-Aza	30	21*	12	0

Table 19: Maternal crosses of NYR-v (NYR-v x WT) displaying endosperm YFP phenotype

Cross number	Active	Silenced	Total count	Percentage active
1	12	50	62	19.4%
2	47	45	92	51.0%
3	32	44	76	42.1%
4	7	8	15	46.7%
5	80	112	192	41.7%
6	43	44	87	49.4%
7	1	18	19	5.3%
8	21	91	112	18.8%
Total	243	412	655	37.1%

Table 20: Paternal crosses of NYR-v (WT x NYR-v) displaying endosperm YFP phenotype

Cross number	Active	Silenced	Total count	Percentage active
1	0	49	49	0.0%
2	5	70	75	6.7%
3	0	31	31	0.0%
4	0	22	22	0.0%
5	0	31	31	0.0%
6	0	63	63	0.0%
7	0	84	84	0.0%
8	0	48	48	0.0%
Total	5	398	403	1.2%

Figure 70: Detailed schematic diagram of NYR-v transgene. LTR, Long-Terminal Repeat; UPN, Upstream NPTII; TR, tandem repeat; HSP70, Heat Shock Protein 70; H2B5, Histone 2B5; YFP, Yellow Fluorescence Protein; e35S CaMV, enhanced 35S Cauliflower Mosaic Virus promoter. Positions relative to e35S transcriptional start site (+1).

ATAGAACCTAAGTGGTTTTGGTGCCATCTTCGAGGTGAACACGGTTGGTCTTTGGATTAGGGTTAGAATTTGACTTAGAGAGAAATAGAGAGCTCTAGCTCTCTCCACTTGGAGATTTGAATAGGGGTTTCTCCAGGGGTCTTTTGC
gypsy-like LTR
AATATTTCCCTCCCCTAATTGCTGGTTATAAGGTTACTTAAGGTTTGGGTAAATGTTACTCATGCTTTGCACACTTGTTCAATCTCTTTTCCCTCTCATCAAGGGTTTTGGGTGATAGGGTTTAAGAGAGGTCTAGGGTTCTTCTCCCCT
gypsy-like LTR
CTTATCCAAGGTAGTAAGTGTGCAAGGATTAGAGTAATGCTTTTTGGTCACTCACACAATCCTTGGCTAAAATTTAGCTTTCCAAGGCCCTCATGCTTTTGCTCCCTTAAGTTCTTCCACATGTGATCATAGCCTTTAGTATCAAGGTGTGCTC
gypsy-like LTR
TAGCCATGGTAACACCTGGGGTGTTACAGCCCCCTTTGCCGGACCTATCTCGCCGCCGGTAAGCCGCGCCACCCTTTCTTTCCCGCGGGTACTGTTTACATTAGGGGAAGGACCTCGGGGAAGAAGGAAAGGAAGGTCAGGGGGCTTTTTG
gypsy-like LTR
AACTGTCAGCGACCCAGGGGAATAATGGCGCAGGGGTAGATTTGTGCGGGTTGATTTTAGTTTTAAACCCAGGGGCCTCGGTGTAAACTGATTTTCCAGGAAACTCTTATTAAATCTATTTTTAAATTGAACATAAACTTTAACAATTCATA
gypsy-like LTR
ACTTCGGTTTAGTTCATCCAAATTTGGTCAAACCAATTTTGCCAGACTTTAAATAATGTAAACTACTTAAGAAAAATATAAAAACCCCTATGTACTACACAGAAAACCTTAGGGTTCTAATTTAATTATTGTTTTGGCCATAAGCTAAATAA
gypsy-like LTR
TAGGGCAAAAAATTAGTTGCACTATTTGACTAGGGTTAGAAAAATTTGGAGAAGTTTATATCTACCTACTGATCTGTTTAAAAATGTTAAACCTCTCTGTCCACCCCATTAAGTTAGTTTTTGCCCCTTATTTGACAATATGCTTAAGGT
gypsy-like LTR
TAAGAGTCCTGCAGTTCATTCAGGGCACCGGACAGGTCGGTCTTGACAAAAAGAACCGGGCGCCCCTGCGCTGACAGCCGGAACACGGCGGCATCAGAGCAGCCGATTGTCTGTTGTGCCCAGTCATAGCCGAATAGCCTCTCCACCCAA
gyp...TR1
NPTII (UPN)
GCGGCCGGAGAACCTGCGTGCAATCCATCTTGTTCAATCATGCGAAACGATCCTCATCCTGTCTCTTGATCAGATCTTGATCCCCTGCGCCATCAGATCCTTGGCGGCAAGAAAGCCATCCAGTTTACTTTGCAGGGCTTCCCAACCTTA
NPTII (UPN)
CCAGAGGGCGCCCCAGCTGGCAATTCCGGTTCGCTTGCTGTCCATAAAACCGCCCAGTCTAGCTATCGCCATGTAAGCCCACTGCAAGCTACCTGCTTTCTCTTTGCGCTTGCGTTTTCCCTTGTCCAGATAGCCCAGTAGCTGACATTC
NPTII (UPN)
ATCCGGGGTCAGCACCGTTTCTGCGGACTGGCTTTCTACGTGTTCCGCTTCCTTTAGCAGCCCTTGCGCCCTGAGTGCTTGCGGCAGCGTGAAGCTGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGG
NPTII (UPN)
CGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTGCGGCCGCGGTTAACAAGCTTCTGCAGGTCCGATTGAGACTTTTCAACAAAGGGTAATATCCG
e35s TR1
GAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTCCGCCGACAGTGGTCCCAAAGATGGAC
e35s TR1
e35s TR2
CCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATGGTCCGATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACT
e35s TR2
TTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTCCGCCGACAGTGGTCCCAAAGATGGACCCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTC
e35s TR2

+1
e35S TR2
e35S unique region
e3_n
HSP70i
HSP70i
HSP70i
HSP70i
HSP70i
HSP70i
H2B5
H2B5
H2B5
H2B5
YFP
YFP
YFP
YFP
YFP
YFP
Nos terminator

Figure 71: Nucleotide sequence of NYR-v transgene. Sequence as determined from genome-walking analysis.

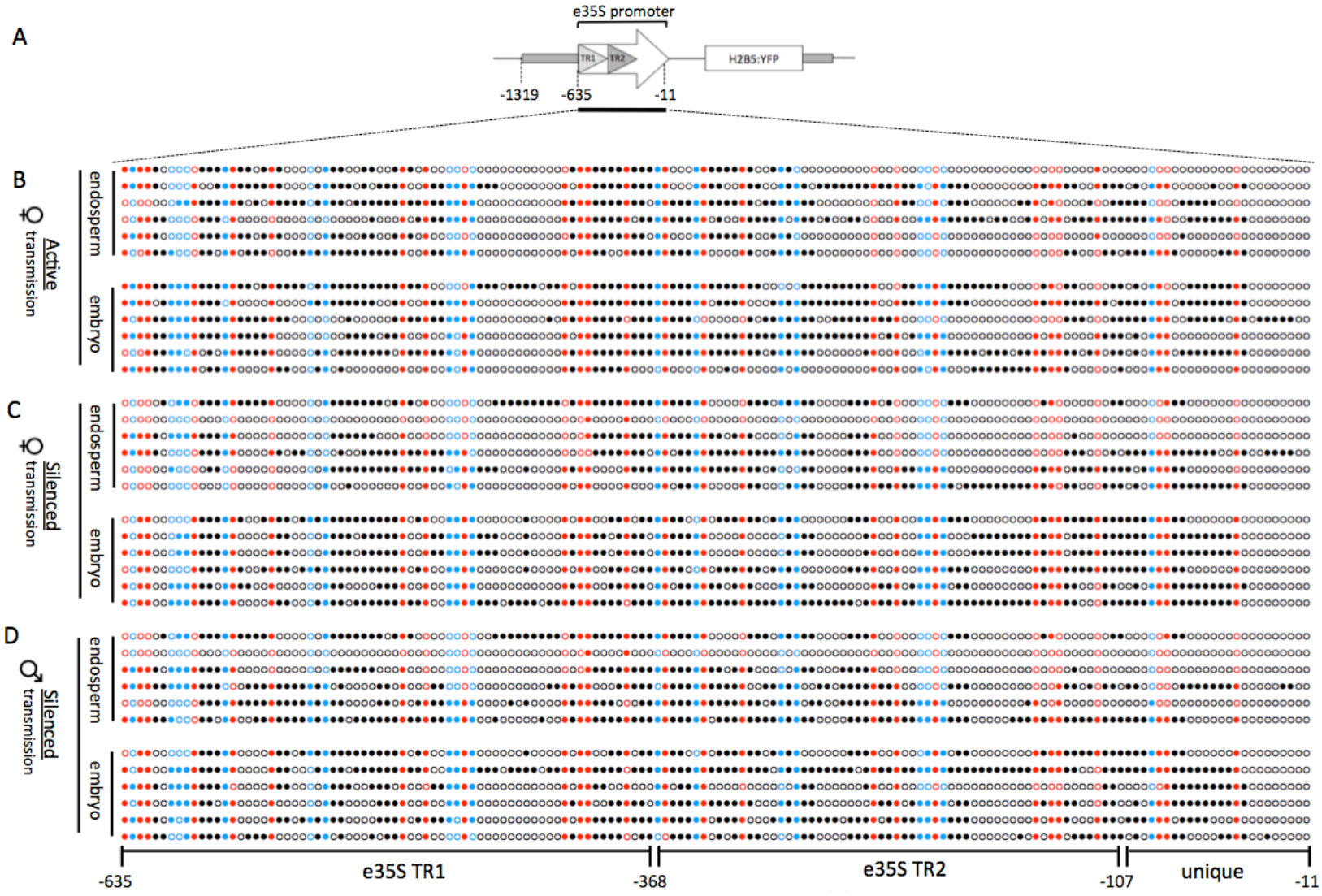


Figure 72: Methylation status of cytosines in the e35S promoter region of the NYR-v following different transmission orientations. (A) Schematic of NYR-v transgene and analysed location in (B) maternally transmitted NYR (active), (C) maternally transmitted NYR-v (silenced) and (D) paternally transmitted NYR endosperm and embryo tissues. The lollipop diagram displays unmethylated cytosines (open circles), methylated cytosines (filled circles), site of asymmetric CHH methylation (black), symmetric CG methylation (red) or symmetric CHG methylation (blue) of 6 clones from bisulfite sequencing.

Table 21: Bisulfite sequencing methylation data of the e35S promoter in endosperm and embryo of maternally and paternally transmitted NYR-v.

Transmission	No. of seeds	Tissue	YFP Expression	% CG	% CHG	% CHH	% Total Methylation
♀	25	endosperm	active	69.23	54.63	41.21	47.51
		embryo	silenced	89.10	75.93	51.97	61.04
♀	25	endosperm	silenced	50.64	41.67	35.45	38.74
		embryo	silenced	96.79	79.63	52.73	63.31
♂	25	endosperm	silenced	53.21	50.93	40.61	43.94
		embryo	silenced	90.38	75.93	48.39	58.78

Transmission	Tissue	YFP Expression	% unique 3' region	% CAAT1-like	% CAAT2-like	% CAAT3-like	% *asf-1* region	% CG (-78)	% CG (-66)
♀	endosperm	active	27.8	55.6	33.3	37.5	37.5	50.0	33.3
	embryo	silenced	45.8	72.2	33.3	66.7	70.8	100.0	83.3
♀	endosperm	silenced	29.2	38.9	66.7	33.3	58.3	50.0	66.7
	embryo	silenced	53.5	94.4	100.0	66.7	100.0	100.0	100.0
♂	endosperm	silenced	17.4	27.8	50.0	16.7	29.2	16.7	33.3
	embryo	silenced	49.3	72.2	100.0	62.5	87.5	83.3	100.0

Table 22: Total size of small RNA sequencing dataset.

Total reads	29,052,106
Reads identified with a single barcode (NYR-v)	10,494,510
Reads identified with a single barcode (B73 WT)	9,447,147
Reads with no barcode	2,758,911
Reads identified as adaptor dimers	7,018,736

7.3 Appendices for Chapter 4

Table 23: Abundance of small RNA species in NYR-v and WT control libraries. 15-29 nt small RNAs in (A) NYR-v and (B) WT control. RPM, Reads per Million

A

Size (nt)	Total Normalised Frequency (RPM)	Total Frequency (Reads)	Unique Frequency (Reads)
15	2733.43	28686	2170
16	3435.80	36057	2375
17	4328.93	45430	2806
18	6938.01	72811	3351
19	7733.19	81156	4101
20	13203.95	138569	5241
21	44333.47	465258	11872
22	75445.35	791762	28792
23	34836.98	365597	19105
24	719726.22	7553174	278406
25	32121.18	337096	31824
26	7940.72	83334	4850
27	6060.31	63600	3109
28	4922.57	51660	2434
29	82.23	863	226

B

Size (nt)	Total Normalised Frequency (RPM)	Total Frequency (Reads)	Unique Frequency (Reads)
15	4823.20	45531	6772
16	7983.37	75363	8997
17	12570.55	118666	12559
18	21848.09	206246	17132
19	28452.12	268588	22301
20	57223.09	540186	27881
21	52643.54	496955	45264
22	118170.87	1115533	82375
23	56003.18	528670	73741
24	459717.90	4339737	672489
25	50197.56	473865	80858
26	31409.85	296509	29549
27	28982.52	273595	29200
28	24897.67	235034	21127
29	438.14	4136	1191

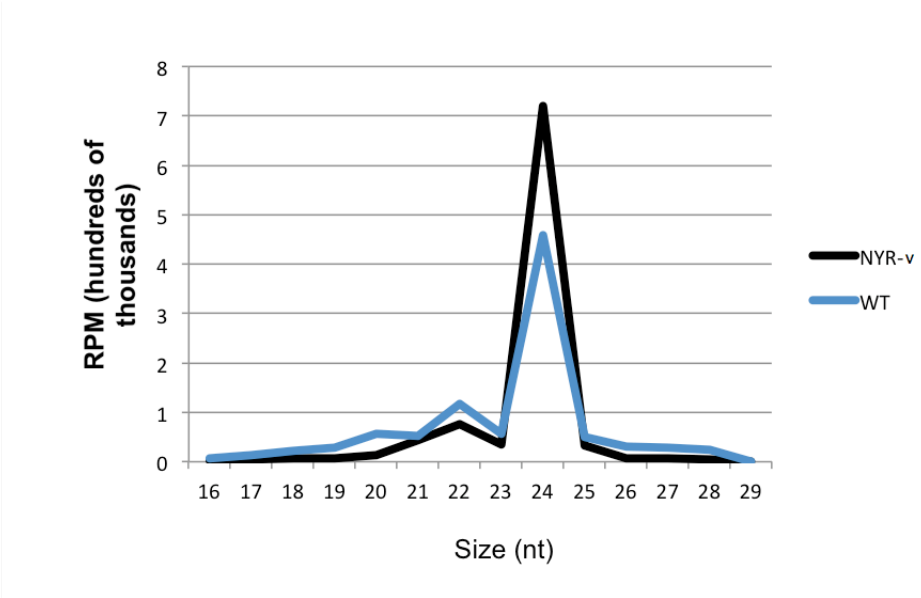


Figure 73: Small RNA profiles of WT and NYR-v libraries. Abundance of small RNA sequences detected in WT and NYR-v libraries between 16-29 nt. RPM, Reads per Million.

Table 24: Distribution and abundance of small RNA species across the NYR-v transgene. Small RNA sequences aligning to NYR-v split between different regions of the transgene. RPM, Reads per Million.

Region	Position	RPM
5' region	-1319 to -637	0
e35S promoter	-636 to +29	58.14
HSP70	+30 to +848	0
H2B	+849 to +1292	14.04
YFP	+1293 to +2020	0
NOS terminator	+2021 to +2301	5.44
Lac	+2302 to 3097	19.18
Delta-kan	+3098 to 3880	15.53
Delta-kan HSP70	+3881 to +4022	0

Table 25: Levels of cytosine methylation across different sequence contexts of the NYR-v e35S promoter in WT and RdDM mutant backgrounds.

Genotype	YFP Expression	% CG	% CHG	% CHH	% Total Methylation
WT	variegated	87.69	85.56	67.45	72.99
(I_4) *mop1-1/mop1-1*	active	28.97	36.30	24.42	26.58
(I_4) *rmr1-1/rmr1-1*	active	44.62	38.89	28.91	32.72
(I_4) *Mop2-1/Mop2-1*	active	35.12	62.59	26.18	31.95
(I_2) *rmr6-1/rmr6-1*	active	44.62	45.19	35.09	37.88

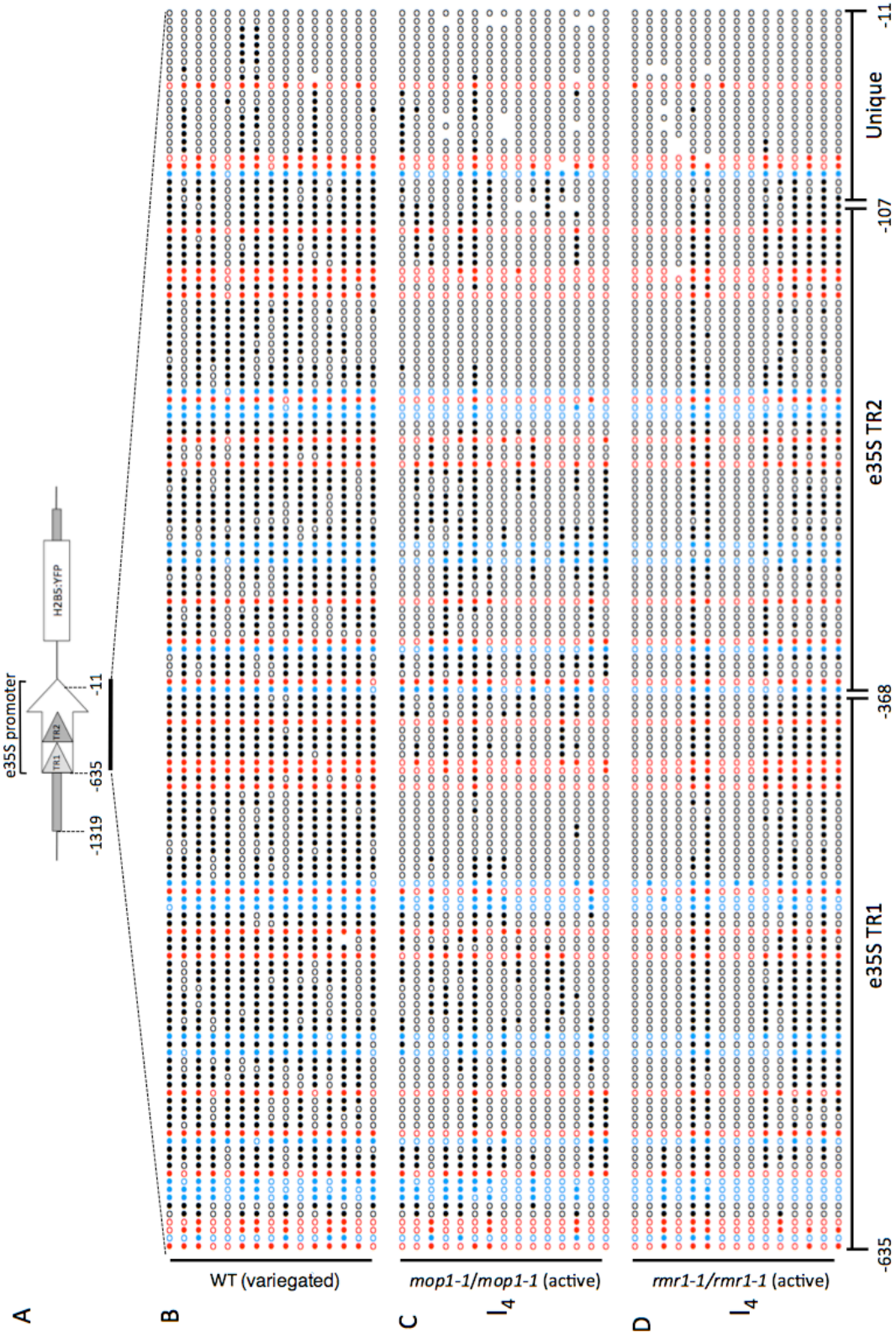


Figure 74: Methylation status of cytosines in the e35S promoter region of NYR in WT and RdDM mutant backgrounds. (A) Schematic of NYR-v transgene and analysed location in (B-D) WT (silenced), I_4 *mop1-1/mop1-1* (active) and I_4 *rmr1-1/rmr1-1* (active) backgrounds. The lollipop diagram displays unmethylated cytosines (open circles), methylated cytosines (filled circles), site of asymmetric CHH methylation (black) and symmetric CG (red) or CHG methylation (blue) of 15 clones from bisulfite sequencing.

Figure 75: Methylation status of cytosines in the e35S promoter region of the NYR in additional RdDM mutant backgrounds. (A) Schematic of NYR-v transgene and analysed location in (B-C) I_4 *Mop2-1/Mop2-1* (active) and I_2 *rmr6-1/rmr6-1* (active) backgrounds. The lollipop diagram displays unmethylated cytosines (open circles), methylated cytosines (filled circles), site of asymmetric CHH methylation (black) and symmetric CG (red) or CHG methylation (blue) of 15 clones from bisulfite sequencing.

7.4 Appendices for Chapter 5

Table 26: Levels of cytosine methylation across the BTG promoter with and without the presence of NYR-v.

Genotype	BTG Expression	% CG	% CHG	% CHH	% Total Methylation	% unique 3' region	% CAAT1-like	% CAAT2-like	% CAAT3-like	% asf-1 region	% CG (-78)	% CG (-66)
BTG-a/-	active	2.0	0.0	0.7	0.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0
NYR-v/-; BTG-v/-	variegated-strong	79.3	70.0	47.7	54.0	56.4	70.0	100.0	72.5	80.0	100.0	100.0
NYR-v/-; BTG-v/-	variegated-weak	2.0	0.0	0.6	0.7	0.4	33.3	0.0	0.0	0.0	0.0	0.0
NYR-v/-; BTG-s/-	silenced	78.0	68.9	45.1	51.7	54.4	70.0	100.0	70.0	85.0	100.0	100.0
-/-; BTG-s/-	silenced	74.7	75.6	42.7	49.9	64.0	100.0	90.0	90.0	97.5	100.0	100.0

Table 27: Levels of cytosine methylation across the BTG promoter in plants derived from endosperm and embryo grafting.

Genotype	BTG Expression	% CG	% CHG	% CHH	% Total Methylation	% unique 3' region	% CAAT1-like	% CAAT2-like	% CAAT3-like	% asf-1 region	% CG (-78)	% CG (-66)
BTG-a/- control	active	3.0	0.0	0.8	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0
BTG-v/- (G6)	variegated	2.0	1.1	1.0	1.1	0.8	0.0	10.0	0.0	0.0	0.0	0.0
BTG-v/- (G8)	variegated	0.67	0.0	1.6	1.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table 28: Progeny derived from outcrossing NYR-v, BTG-v plants.

Genotype	BTG expression	Expected freq.	Observed number	Observed freq.
NYR-v; BTG-s/-	silent	37.5%	7	35.00%
NYR-v; -/-	n/a	37.5%	8	40.00%
-/-; BTG-s/-	silent	12.5%	2	10.00%
-/-; -/-	n/a	12.5%	3	15.00%
		Total	20	

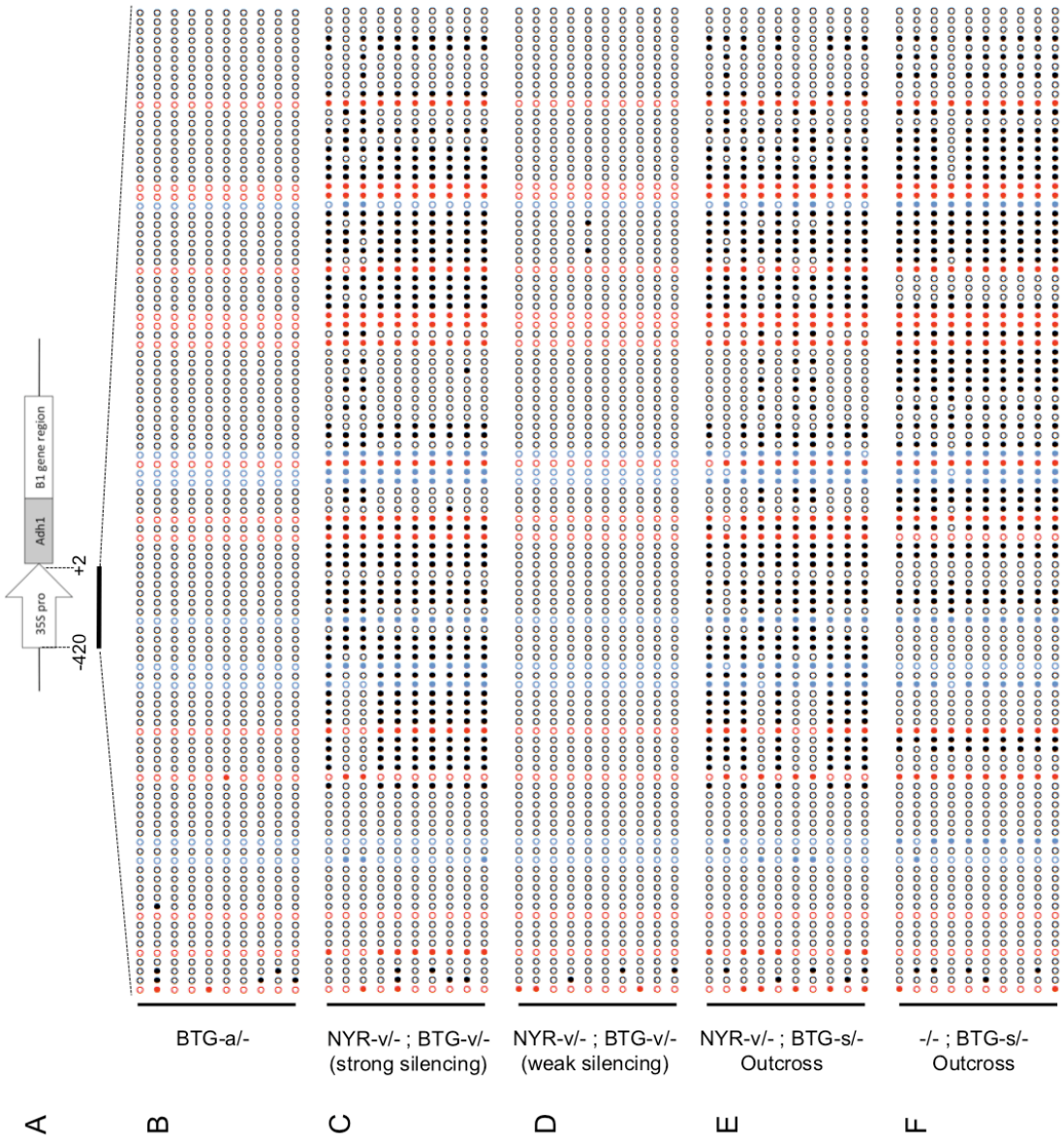


Figure 76: Methylation status of cytosines in the 35S promoter region of the BTG with and without NYR-v. (A) Schematic of BTG transgene and analysed location in (B-F) BTG-a/- (no NYR-v contact), NYR-v/-; BTG-v/- (strong silencing), NYR-v/-; BTG-v/- (weak silencing), NYR-v/-; BTG-s/- (outcross) and -/-; BTG-s/- (outcross). The lollipop diagram displays unmethylated cytosines (open circles), methylated cytosines (filled circles), site of asymmetric CHH methylation (black) and symmetric CG (red) or CHG methylation (blue) of 10 clones from bisulfite sequencing.

Figure 77: Methylation status of cytosines in the 35S promoter region of the BTG with and without NYR-v. (A) Schematic of BTG transgene and analysed location in (B-D) BTG-a/- derived from a BTG endosperm and embryo graft and grafted plants G6 and G8 derived from growth with NYR endosperm showing varieagted colour expression. The lollipop diagram displays unmethylated cytosines (open circles), methylated cytosines (filled circles), site of asymmetric CHH methylation (black) and symmetric CG (red) or CHG methylation (blue) of 10 clones from bisulfite sequencing.

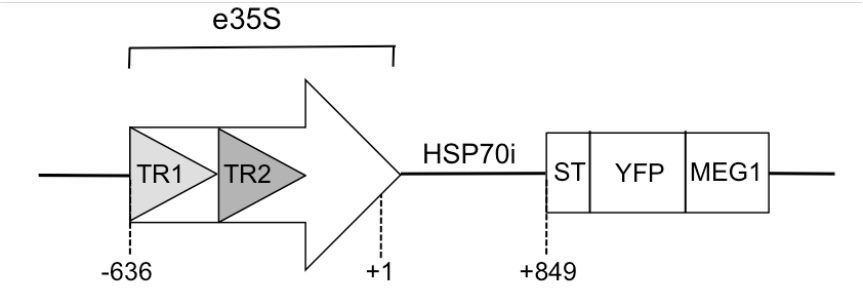


Figure 78: Schematic diagram of Cell wall YFP Reporter (CYR) transgene. TR, tandem repeat; e35S, enhanced 35S CaMV promoter; HSP70i, Heat shock protein 70 intro; ST, signal secretion peptide; YFP, yellow fluorescent protein; MEG1, maternally expressed gene 1. Positions relative to e35S transcriptional start site (+1).

Table 29: Progeny of NYR-v and CYR-a crosses. Three cross events conducted between hemizygous NYR-v and homozygous CYR-a transgenics

Cross number	Genotype	Number
1	-/-; CYR-a/-	13
	NYR-v/-; CYR-v/-	17
		30
2	-/-; CYR-a/-	15
	NYR-v/-; CYR-v/-	15
		30
3	-/-; CYR-a/-	14
	NYR-v/-; CYR-v/-	16
		30